Models of spatial representation in the medial entorhinal cortex

The origin, inheritance, and amplification of grid-cell activity

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

eingereicht an der
Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von

M.Sc. Tiziano D'Albis

Präsidentin der der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät:
Prof. Dr. Bernhard Grimm

Gutachter/innen:

1. Prof. Dr. Richard Kempter
2. Prof. Dr. Michael Brecht
3. Prof. Dr. Alessandro Treves

Tag der mündlichen Prüfung: 12. Juli 2018

Acknowledgments

This thesis is the fruit of five years of work under the supervision of Prof. Richard Kempter at the institute for theoretical biology in Berlin. Here, I spent an amazing period of my life, which has passed incredibly fast. This is thank to the very friendly, helpful, and intellectually stimulating environment that have welcomed me since the first day. I wish I could remain much longer.

I shall thank above all Richard, my supervisor, who had supported me with great enthusiasm since my very first steps in neuroscience. Without Richard, this would have been certainly a different PhD.

I am indebted to my officemates Eric and Nathalie for the relaxed, fun, and very supportive atmosphere we could create in room 013. They also contributed immensely to this work by proof-reading the text, checking the math, and helping out with the German translation of this thesis abstract. I shall also thank André, Pia, and Paul for providing valuable feedback on the text, and Martina for guidance on the thesis structure.

This work would have not been possible without the support of all my colleagues and friends who remained close throughout this journey. Thanks to Roberta, Susana, José, Jorge, and Nikolay for providing a solid basis to the late-lunch group—always a source of great inspiration for both research and non-research matters. Also thanks to Thomas, Jannina, Paula, Jan-Hendrik, Michiel, Paul, Katha, Fred, and Susanne for stimulating scientific discussions.

Many other people contributed to create an incredible research environment much beyond the walls of our institute. I shall mention Michael, Nacho, Rajnish, Shimpei, and Jean in the Brecht lab, and Dietmar, Jochen, Prateep, Nikolaus, and Barbara in the Schmitzlab. And of course a big thank to the BCCN Berlin, where my excitement for neuroscience blossomed and will certainly remain flourished in the years to come.

Tiziano D'Albis
Berlin, April 2018

Abstract

High-level cognitive abilities such as memory, navigation, and decision making rely on the communication between the hippocampal formation and the neocortex. At the interface between these two brain regions is the entorhinal cortex, a multimodal association area where neurons with remarkable representations of self-location have been discovered: the grid cells.

Grid cells are neurons that fire according to the position of an animal in its environment. A single grid cell activates at multiple spatial locations with firing fields that are arranged in a strikingly-regular triangular pattern. Grid cells are thought to support animal's navigation and spatial memory, but the cellular mechanisms that generate their patterns are still unknown. In this thesis, I study computational models of neural circuits to explain the emergence, inheritance, and amplification of grid-cell patterns.

In the first part of the thesis, I focus on the initial formation of grid firing fields. I embrace the idea that periodic representations of space could emerge via a competition between persistently-active spatial inputs and the reluctance of a neuron to fire for long stretches of time. Building upon previous theoretical work, I propose a single-cell model that generates grid-like activity solely form spatially-irregular inputs, spike-rate adaptation, and Hebbian synaptic plasticity. Compared to previous proposals, my model achieves a higher level of biological realism, gives unprecedented analytical insights, and generates novel experimental predictions.

In the second part of the thesis, I focus on the inheritance and amplification of grid-cell patterns. Motivated by the architecture of entorhinal microcircuits, I investigate how feed-forward and recurrent connections affect grid-cell tuning. I show that grids can be inherited across neuronal populations, and that both feed-forward and recurrent connections can improve the regularity of spatial firing. Finally, I show that a connectivity supporting these functions could self-organize in an unsupervised manner.

Altogether, this thesis contributes to a better understanding of the principles governing the neuronal representation of space in the medial entorhinal cortex.

Zusammenfassung

Komplexe kognitive Funktionen wie Gedächtnisbildung, Navigation und Entscheidungsprozesse hängen von der Kommunikation zwischen Hippocampus und Neokortex ab. An der Schnittstelle dieser beiden Gehirnregionen liegt der entorhinale Kortex—ein Areal, das Inputs verschiedener Sinne assoziiert und das Neurone mit bemerkenswerten räumlichen Repräsentationen enthält: Gitterzellen.

Gitterzellen sind Neurone, die abhängig von der Position eines Tieres in seiner Umgebung feuern. Eine einzelne Gitterzelle ist an mehreren Orten im Raum aktiv und bildet Feuerfelder in einem auffallend dreieckigen Muster. Man vermutet, dass Gitterzellen Navigation und räumliches Gedächtnis unterstützen, aber die zellulären Mechanismen, die diese Muster erzeugen, sind noch immer unbekannt. In dieser Dissertation untersuche ich mathematische Modelle neuronaler Schaltkreise, um die Entstehung, Weitervererbung und Verstärkung von Gitterzellaktivität zu erklären.

Im ersten Teil dieser Dissertation konzentriere ich mich auf die ursprüngliche Entstehung von Gittermustern. Ich greife die Idee auf, dass periodische Repräsentationen des Raumes durch Konkurrenz zwischen dauerhaft aktiven, räumlichen Inputs und der Tendenz eines Neurons durchgängiges Feuern zu vermeiden entstehen könnten. Aufbauend auf vorangegangenen theoretischen Arbeiten stelle ich ein Einzelzell-Modell vor, das gitterartige Aktivität allein durch räumlich-irreguläre Inputs, Feuerratenadaptation und Hebbsche synaptische Plastizität erzeugt. Verglichen mit bisherigen Ansätzen erreicht mein Modell eine größere biologische Plausibilität, ermöglicht nie da gewesene analytische Einsichten und erbringt neue, experimentelle Vorhersagen.

Im zweiten Teil der Dissertation liegt der Schwerpunkt auf der Vererbung und Verstärkung von Gitterzellmustern. Motiviert durch entorhinale Mikroschaltkreise untersuche ich den Einfluss von vorwärts gerichteten und rekurrenten Verbindungen auf das Gitter-Tuning. Ich zeige, dass Gittermuster zwischen neuronalen Populationen weitervererbt werden können und dass sowohl vorwärts gerichtete als auch rekurrente Verbindungen die Regelmäßigkeit von räumlichen Feuermustern verbessern können. Schließlich zeige ich, dass eine entsprechende Konnektivität, die diese Funktionen unterstützt, autonom, auf unüberwachte Weise entstehen könnte.

Insgesamt trägt diese Arbeit zu einem besseren Verständnis der Prinzipien der neuronalen Repräsentation des Raumes im medialen entorhinalen Kortex bei.

Contents

List of abbreviations

5HT3aR serotonin-3a receptor. 15, 16

AHP after-spike hyperpolarizing potential. 17, 74, 87–89

BCM Bienenstock-Cooper-Munro. 45

CA cornu ammonis. 7–9, 13, 19, 74, 152

CAN continuous attractor network. 29–35, 40, 41, 103, 105, 107, 120–122, 125, 143, 150

CCK cholecystokinin. 14, 16

DG dentate gyrus. 8, 13

EC entorhinal cortex. 8–10, 15

HCN hyperpolarization-activated cyclic nucleotide-gated. 74, 75, 142

HD head direction. 26, 27

$\mathbf{I_h}$ hyperpolarization-activated cation current. 16, 17, 74, 75

LEC lateral entorhinal cortex. 8–10

LTD long-term depression. 45, 75, 77

LTP long-term potentiation. 45, 75, 77

MEC medial entorhinal cortex. 8–10, 12–17, 19–25, 27, 36, 39, 45–47, 51, 52, 68, 73–78, 95–97, 99, 103, 116–122, 141–144, 150

NMDA N-methyl-D-aspartate. 76, 77, 142, 149

OI oscillatory interference. 36, 39–41

PCA principal component analysis. 47

PER perirhinal cortex. 8–10

POR postrhinal cortex. 8–10

PV parvalbumin. 14–16

SOM somatostatin. 14–16

STDP spike-timing-dependent plasticity. 55, 58, 59, 65, 66, 75

VCO velocity-controlled oscillator. 37–40

Chapter 1

Introduction

When we visit a new environment—such as a foreign city—we are initially disoriented. To find our way to a museum or a local restaurant, we need to rely on printed maps picked up at the tourist kiosk or electronic maps on our smartphones. Yet, as the environment becomes more familiar, we are soon able to recognize previously encountered routes and landmarks and even get back to our hotel taking detours or shortcuts. But how do we learn spatial concepts? And how is physical space represented in the brain?

A prominent theory in experimental psychology posits that men and other animals can interiorize spatial experiences in a sort of 'cognitive map' of the environment, i.e., a mental representation of space that embeds known places and their relations into a common reference frame (Tolman, 1948). For example, when a friend asks for directions to our home, we can create a mental image of the roads, turning points, and landmarks along the way. This representation is a cognitive map.

The cognitive-map theory was introduced by the psychologist Edward Tolman to explain rodents' spatial behavior (Tolman, 1948). Tolman himself cites an anecdote reported by Karl Lashley in 1929. Lashley once discovered that some rats, after having learned a maze, they "pushed back the cover near the starting box, climbed out and run directly across the top to the goal box where they climbed down again and ate" (Tolman, 1948). Such a behavior was in striking contrast to the common belief at the time that rodents simply use stimulus-response associations to solve mazes—alike to learning the right connections in a complicate telephone switchboard. In fact, through a series of ingenious experiments, Tolman (1948) demonstrated that rats' spatial behavior resembled more a sophisticated control room than an old-fashioned telephone exchange. Tolman writes: "in the course of learning something like a field map of the environment gets established in the rat's brain" (Tolman, 1948).

Tolman's cognitive map, however, remained a purely psychological concept for a long time. It was not until O'Keefe and Dostrovsky (1971) discovered hippocampal place cells that clear correlates of spatial behavior were uncovered in the brain. By recording from the dorsal hippocampus of freely-moving rats, O'Keefe and Dostrovsky

(1971) found neurons that responded "solely or maximally when the rat was situated in a particular part of the testing platform facing in a particular direction". These neurons—which are now called place cells—were immediately seen by the experimenters as the neural substrate of the cognitive map theorized by Tolman more than twenty years before.

The discovery of hippocampal place cells sparked an enormous interest in the field of systems neuroscience, because, for the first time, a high-level cognitive concept—such as the perception of one's location in the environment—could be studied at a mechanistic level, bridging a large gap between psychology and physiology.

Thirty-four years after the remarkable report by O'Keefe and Dostrovsky (1971), another major breakthrough was made in the field of systems neuroscience. In a quest of unveiling the inputs to hippocampal place cells, Hafting et al. (2005) discovered neurons in the medial entorhinal cortex that fired at multiple locations in the environment, forming an strikingly-regular triangular pattern. These neurons were termed grid cells.

Entorhinal grid cells represent physical space with periodic tuning curves at multiple spatial scales, a remarkably efficient code which was shown to be optimal in information theoretical terms (Wei et al., 2015; Stemmler et al., 2015)—and which no theorist had predicted to be actually implemented in the brain. After more than ten years of investigation, however, it remains to date unclear how grid patterns emerge and how grid-cell activity is processed within the cortex.

1.1. Aim and scope of the thesis

The aim of the this thesis is to shed light on the neuronal mechanisms underpinning the origin, inheritance, and amplification of grid cell-activity in the medial entorhinal cortex. To this aim, I study mathematical models of neural circuits explaining currently available empirical data and making quantitative predictions for future experiments.

1.2. Outline of the thesis

The thesis is structured as follows. In Part I, I provide the reader with an adequate background to understand the original research presented afterwards. Here, I critically review both experimental and theoretical studies on grid-cell activity and the brain regions where it is found. I start by giving a broad overview to the hippocampal system and the entorhinal cortex, with a focus on the anatomy of the tissue and the physiology of the neurons (Chapter 2). I then summarize the main empirical findings on grid-cell patterns, including their geometrical properties, modular organization, and influence by sensory cues (Chapter 3). Finally, I review the main theories on

grid-pattern formation, confronting the models with empirical evidence, and pointing out the experimental findings that remain to be explained (Chapter 4).

In Part II, I present the original research work of this thesis. First, I propose a model for the emergence of grid-cell activity (Chapter 5). Through analytical work and simulations, I show that grid-like patterns can emerge from a single-cell learning process based on synaptic plasticity and neuronal adaptation. Next, I study how grid-cell activity may be affected by network dynamics (Chapter 6). I show that grid patterns can be inherited by feed-forward projections and even amplified by both feed-forward and recurrent circuits. Finally, I study how grid-cell microcircuits may develop during ontogeny, and I outline a coherent framework in which both external (sensory) and internal (self-motion) cues could maintain grid-cell activity during exploration.

In Part III, I summarize the results of the thesis, review experimental predictions, and provide directions for future work (Chapter 7). At last, I discuss open questions related to the role of grid-cell activity in high-level cognitive tasks, such as navigation, memory, and abstract cognition (Chapter 8).

Part I.

Literature review

Chapter 2

The hippocampal system and the entorhinal cortex

Here, I review anatomical and physiological data that is relevant to understand grid-cell activity. I start by introducing the reader to the hippocampal system—the brain area where grid cells are found (Section 2.1). In particular, I describe the anatomical architecture of the hippocampal system and its main functions. Next, I focus on the medial entorhinal cortex—the subregion where grid cells are most abundant (Section 2.2). Entorhinal neurons are described in terms of their morphological and physiological traits, identifying local cellular microcircuits that could support grid patterns.

2.1. The hippocampal system

The hippocampal system is a compound cortical structure located in the medial temporal lobe of the mammalian brain (Figure 2.1A). According the nomenclature proposed by Burwell and Agster (2008), it can be subdivided in two main areas: the hippocampal formation (dentate gyrus, hippocampus proper, and subiculum) and the parahippocampal region (entorhinal, perirhinal, postrhinal, presubicular and parasubicular cortices, Figure 2.1B). These two areas are distinguished by the cytoarchitectonic organization of the cortical tissue: the hippocampal formation (allocortex) is characterized by a trilaminar structure and by largely unidirectional internal pathways (Cajal, 1893), whereas the parahippocampal region (periallocortex) is characterized by a six-layered architecture with overwhelmingly reciprocal synaptic connections (Insausti et al., 2017). Within the hippocampal formation, the hippocampus proper—also called the *cornu ammonis* (CA)—is further subdivided in three major anatomical subfields (CA1-3, Lorente de Nó, 1934).

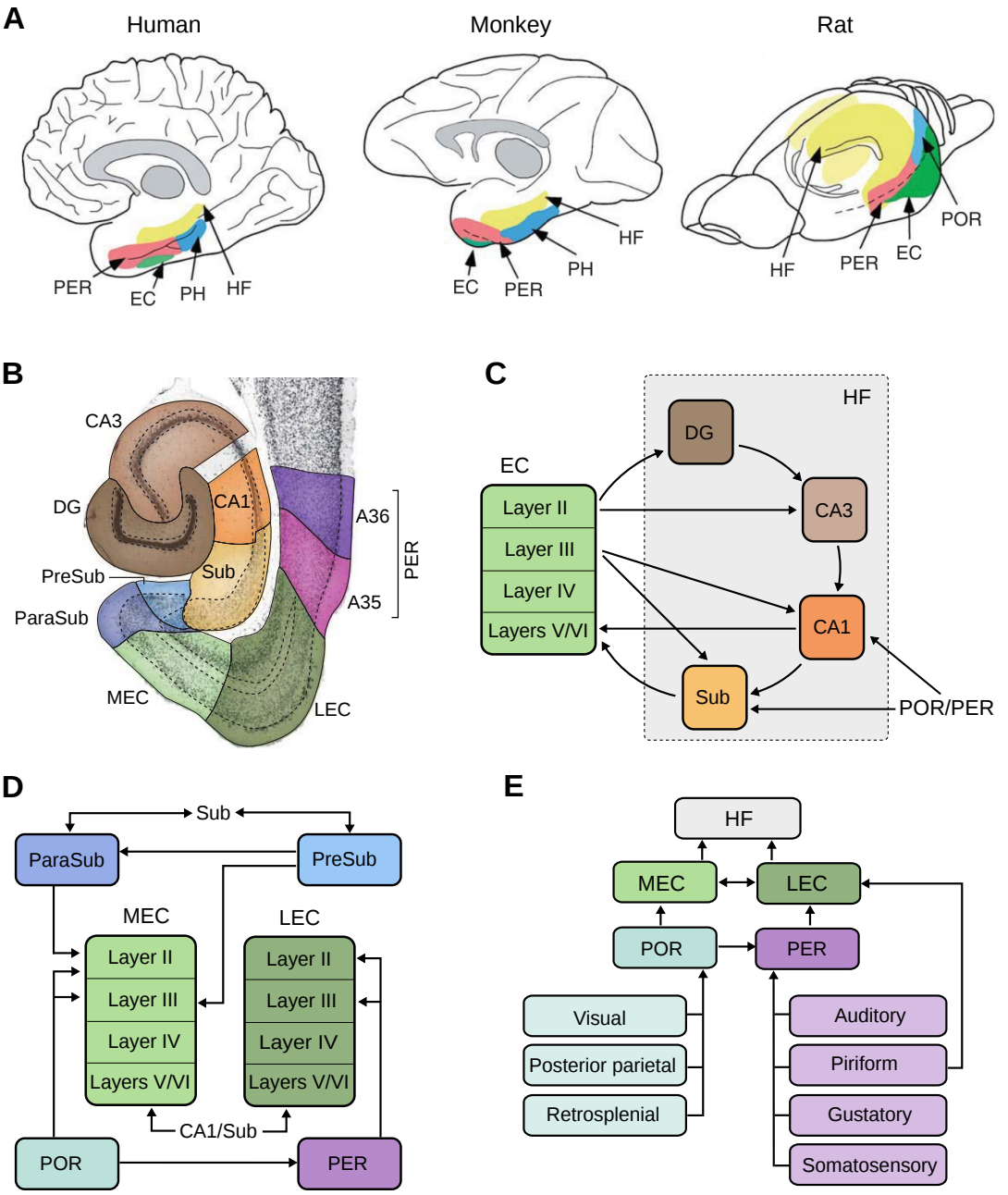

Figure 2.1. | **Anatomy of the hippocampal system**. **A)** The hippocampal system (colored regions) in the human (left), monkey (middle), and rat (right) brain. EC, entorhinal cortex; HF, hippocampal formation; PER, perirhinal cortex; PH, parahippocampal cortex; POR, postrhinal cortex. Note that the primate PH is the homologue of the rodent POR. Image from Burwell and Agster (2008) reproduced with permission from Elsevier. **B)** Nissl-stained horizontal cross section of the rat hippocampal formation and parahippocampal regions. DG, dentate gyrus; CA, *cornu ammonis*; Sub, subiculum; PreSub; presubiculum; ParaSub, parasubiculum; MEC, medial entorhinal cortex; LEC, lateral entorhinal cortex. Brodmann areas 35 and 36 form the perirhinal cortex (PER, see panel A). Image from Van Strien et al. (2009) adapted with permission from Nature Publishing Group. **C)** Simplified diagram of the excitatory connections between the EC and the hippocampal formation. Note that principal cells are nearly absent in layers I and IV. **D)** Simplified diagram of the parahippocampal inputs to the EC. MEC, medial entorhinal cortex; LEC, lateral entorhinal cortex. **E)** Simplified diagram of the neocortical inputs to the EC. Note that only the strongest connections are depicted, see main text for details. Diagrams in panels C-E were drawn based on connectivity data reviewed by Burwell and Agster (2008); Van Strien et al. (2009); Witter et al. (2017).

2.1.1. Anatomy of the hippocampal system

I now describe the main connectivity patterns between components of the hippocampal system. I focus on the anatomy of the rat, which has been studied extensively and is most relevant to grid-cell data. More comprehensive reviews on this topic can be found in (Burwell and Agster, 2008; Van Strien et al., 2009; Witter et al., 2017) and references therein.

The entorhinal-hippocampal circuitry

A key region of the hippocampal system is the entorhinal cortex (EC, Brodmann area 28), which is located at the interface between the hippocampal formation and the neocortex (Figure 2.1B). The EC provides the main input to the hippocampal formation via the *perforant pathway* (Figure 2.1C), i.e., principal cells in EC layer II target the dentate gyrus and the CA3 field, and principal cells in EC layer III target CA1 and the subiculum (*temporoammonic pathway*). Between hippocampal fields, synaptic connections are predominantly feed-forward: from the dentate gyrus to CA3 (mossy fibers), from CA3 to CA1 (Schaffer collaterals), and from CA1 to the subiculum. Finally, the CA1 field and the subiculum project back to the EC by forming synaptic contacts with cells in layer V and VI. Such a connectivity pattern—termed the entorhinal-hippocampal loop—suggests that a largely unidirectional flow of information exists from the superficial entorhinal layers, through the hippocampal formation, and back to the deep entorhinal layers (Andersen et al., 1971; Amaral and Witter, 1989).

The parahippocampal circuitry

The EC receives excitatory input from all parahippocampal regions. The parahippocampal input to the EC is restricted to the superficial layers (layers II and III) and is anatomically organized in two parallel pathways targeting the medial entorhinal cortex (MEC) and the lateral entorhinal cortex (LEC) selectively (Figure 2.1D). Fibers originating from the perirhinal cortex (PER, Brodmann areas 35 and 36) preferentially project to the LEC, whereas those originating from the postrhinal cortex (POR) preferentially target the MEC (Naber et al., 1997). Interestingly, the two entorhinal subfields (MEC and LEC) in turn project to distinct sub-populations of principal cells in the hippocampal formation (within field CA1, and subiculum), suggesting that two parallel streams of information can be processed by the hippocampal system (Naber et al., 1997; Witter et al., 2000). Yet the two pathways are also strongly interconnected: the POR strongly projects to the PER (with weaker return projections), and the two entorhinal subfields form reciprocal synaptic contacts across all layers (Witter et al., 2000).

The MEC is further innervated by both the presubiculum and the parasubiculum.

Although both structures project to all entorhinal layers (Canto et al., 2012), parasubicular axons innervate preferentially MEC layer II, whereas presubicular axons terminate preferentially in MEC layer III (Köhler, 1984; Van Groen and Wyss, 1990; Caballero-Bleda and Witter, 1993, 1994; Preston-Ferrer et al., 2016). The presubiculum projects to the superficial layers of parasubiculum, but the reciprocal connection is nearly absent (Burwell and Agster, 2008). Both subicular cortices are interconnected with the subiculum in the hippocampal formation (O'Mara et al., 2001; Burwell and Agster, 2008).

Neocortical pathways

Widespread cortical input reaches the hippocampal formation via the parahippocampal region (Figure 2.1E). The POR receives mainly visual and visuospatial information from visual, posterior-parietal, and retrosplenial cortices; whereas the PER is recipient of multiple sensory modalities, including inputs from auditory, olfactory, gustatory, and somatosensory cortices (Agster and Burwell, 2009). This segregation of input modalities is consistent with the hypothesis that two functional pathways could coexist: a POR-MEC pathway conveying information about spatial context, and a PER-LEC pathway conveying non-spatial information about sensory stimuli, e.g., object, odors, or auditory stimuli (Burwell, 2000; Witter et al., 2000; Agster and Burwell, 2009).

The EC is also innervated by neocortical areas directly. In particular, the nearby piriform (olfactory) and insular cortices project both to the LEC (strongly) and to the MEC (moderately). The orbitofrontal cortex preferentially targets the LEC, whereas parietal and occipital cortices project to the MEC (Agster and Burwell, 2009). In summary, the MEC constitutes a crucial anatomical hub connecting posterior cortical areas (including parietal and occipital networks) to the hippocampus.

Subcortical pathways

The EC is also heavily interconnected with subcortical structures. Subcortical projections arise in the claustrum, the olfactory bulb, the amygdala, the thalamus, the hypothalamus, the brain stem, and the medial septum (Agster and Burwell, 2009). Septal projections to the EC (and to other areas of the hippocampal system) arise from glutamatergic, GABAergic, and cholinergic neurons (Manns et al., 2001). Specifically, GABAergic and cholinergc projections contribute to the entrainment of the hippocampal theta rhythm (4–12 Hz, Barry et al., 2012b; Gonzalez-Sulser et al., 2014; Colgin, 2016), one of the most regular extracellular oscillations in the rodent brain (Vanderwolf, 1969; Winson, 1974; Buzsáki, 2002). The input from the medial septum is especially relevant to grid-cell activity, because septal inactivations disrupt theta rythmicity and impair spatial memory and grid-cell firing (Mitchell et al., 1982; Brandon et al., 2011;

Koenig et al., 2011).

2.1.2. Functions of the hippocampal system

I now discuss the main behavioral correlates of the hippocampal system. Although hippocampal function is still hotly debated (Lisman et al., 2017), a large body of experimental evidence points to two—partially interrelated—cognitive abilities: episodic memory and spatial cognition (e.g., O'Keefe and Nadel, 1978; Squire and Zola-Morgan, 1991; Redish, 1999; Eichenbaum, 1993; Morris, 2007; Eichenbaum, 2017).

Episodic memory

The hippocampal system has been linked to memory processes since the late 1950s, when Brenda Milner and William Scoville reported the clinical case of patient Henry G. Molaison, also known as H.M. in the literature (Scoville and Milner, 1957). Patient H.M. suffered of intractable temporal-lobe epilepsy and underwent a neurosurgical procedure in which the hippocampal formation, the amygdala, and the entorhinal cortex were removed bilaterally (Corkin et al., 1997; Annese et al., 2014). The operation successfully reduced the incidence of epileptic seizures but, unexpectedly, left the patient with profound memory deficits. In particular, H.M. acquired a severe *anterograde amnesia*, a permanent inability to form new long-term memories about people, places, or objects that were encountered after surgery (Scoville and Milner, 1957; Corkin, 2002).

After patient H.M., a large number of clinical and functional-imaging studies showed that the hippocampal system is important for declarative memory, i.e., the memory of facts and events that can be explicitly declared (Squire, 1992; Burgess et al., 2002; Davachi and Dobbins, 2008; Ekstrom and Ranganath, 2017). Specifically, the hippocamapal system was found to be more important for remembering autobiographical events (episodic memory), rather than factual knowledge (semantic memory; Vargha-Khadem et al., 1997; Tulving and Markowitsch, 1998). In line with this view, episode-like memory traces could be manipulated by directly interfering with hippocampal activity in rodents (see e.g., Roy and Tonegawa, 2017, for a review).

Spatial cognition

The second function commonly attributed to the hippocampal system is spatial cognition. The clearest experimental link between spatial cognition and the hippocampus dates back to the early 1970s, when John O'Keefe and Jonathan Dostrovsky discovered hippocampal place cells, i.e., neurons that selectively activate whenever the animal visits a restricted location of the environment (the place field of the cell, O'Keefe and Dostrovsky, 1971; O'Keefe, 1976). Place cells were immediately seen as the neural substrate of a 'spatial cognitive map', an allocentric internal representation of space that

could support self-location, navigation, and spatial memory (Tolman, 1948; O'Keefe and Nadel, 1978).

Today, we know that neural correlates of space are found in all regions of the hippocampal system through a variety of functional cell types, e.g., head direction cells, grid cells, border cells, speed cells (see, e.g., Hartley et al., 2014, for a review). Additionally, experimental damage to the hippocampus and the parahippocampal region causes severe behavioral deficits in a wide range of spatial-memory tasks, e.g., the radial-arm maze (Olton et al., 1979), the Morris watermaze (Morris et al., 1982), and the T-maze (Rawlins and Olton, 1982). Finally, clinical and neuroimaging studies indicate that the hippocampal system processes spatial information also in humans (e.g., Burgess et al., 2002; Epstein et al., 2017).

Space as a memory trait

The hippocampal system is thus implicated in both episodic memory and spatial cognition, yet these two functions are probably not independent. Every episode we experience has at least a spatial (where) and a temporal (when) component. Because hippocampal research in rodents has largely focused on the neural correlates of space, the spatial component could have been overemphasized. Indeed, several non-spatial variables are equally represented by the hippocampal system, e.g., time (Manns et al., 2007), sounds (Abbott and Blum, 1996; Aronov et al., 2017), odors (Eichenbaum et al., 1987), objects (Fried et al., 1997; Deshmukh and Knierim, 2011), faces (Fried et al., 1997), social bounds (Tavares et al., 2015; Omer et al., 2018), and conceptual knowledge (Constantinescu et al., 2016). By representing any relevant variable in a given cognitive task, the fundamental function of the hippocampal system may be to form associations between those variables, possibly organizing them into temporal sequences (Hasselmo, 2011; Buzsáki and Moser, 2013).

2.2. Architecture of the medial entorhinal cortex

I now focus on the MEC, the sub-region of the hippocampal system where grid cells are most abundant. To understand how grid-cell activity is generated and possibly propagated to different areas, I summarize the main anatomical (Section 2.2.1) and intrinsic physiological properties (Section 2.2.2) of entorhinal neurons.

2.2.1. Microcircuits of the entorhinal cortex

I start by describing cell-type specific excitatory microcircuits, which attracted considerable attention in recent years (see, e.g., Sürmeli et al., 2015; Fuchs et al., 2016; Winterer et al., 2017).

Excitatory cell types

The MEC contains two main excitatory cell types: stellate cells and pyramidal cells (Klink and Alonso, 1997). Stellate cells have multiple primary dendrites that stem from a round soma, whereas pyramidal cells have one prominent apical dendrite arising from an elongated soma (Figure 2.2A). Layer II principal cells have either stellate (65%) or pyramidal morphology (32%), whereas principal cells is layers III, V, and VI are mostly pyramidal[1] (Klink and Alonso, 1997; Canto et al., 2008). Stellate and pyramidal cells are characterized by distinct electrophysiological properties (Section 2.2.2) and by differential immunoreactivity to antibodies, i.e., stellate cells are typically reelin positive (and calibindin negative), whereas pyramidal cells are typically calbindin positive (and reelin negative; Varga et al., 2010; Ray et al., 2014; Kitamura et al., 2014). These distinct immunoreactive responses have been particularly important to unveil cell-type-specific connectivity patterns within the entorhinal circuitry.

Aside from stellate and pyramidal cells, Fuchs et al. (2016) recently identified two additional excitatory cell types in the MEC, i.e., intermediate stellate and intermediate pyramidal cells. This classification, however, was based solely on electrophysiological features, and could not be reproduced by Winterer et al. (2017). Therefore, in the following, I shall focus on experimental data that relates to the two best described cell types: stellate and pyramidal cells.

Excitatory circuits

Layer II stellate and pyramidal cells are characterized by strikingly segregated input and output projection patterns (Figure 2.2C). Layer II stellate cells project to the hippocampal formation (DG and CA3 fields, Varga et al., 2010) and to layer V (Sürmeli et al., 2015), whereas layer II pyramidal cells project to the controlateral MEC (Varga et al., 2010) and to CA1 interneurons (Kitamura et al., 2014). In terms of the inputs, it was found that layer II pyramidal (but not stellate) cells are selectively targeted by the parasubiculum (Tang et al., 2016) and by cholinergic neurons in the medial septum (Ray et al., 2014).

Within layer II, recurrent synaptic excitation was originally thought to be very sparse or absent (Dhillon and Jones, 2000; Couey et al., 2013; Pastoll et al., 2013), but recent studies challenged this hypothesis (Fuchs et al., 2016; Winterer et al., 2017; Schmidt et al., 2017); see also (Kumar and Buckmaster, 2006). Notably, Schmidt et al. (2017) found that a typical layer II principal cell makes more than one third of its synaptic contacts onto excitatory neurons of the same layer, which is consistent with recent reports by Fuchs et al. (2016) and Winterer et al. (2017). Yet recurrent excitation is even stronger within layers III and V (Dhillon and Jones, 2000; Winterer et al., 2017).

[1] Layers I and IV are largely devoid of principal cells.

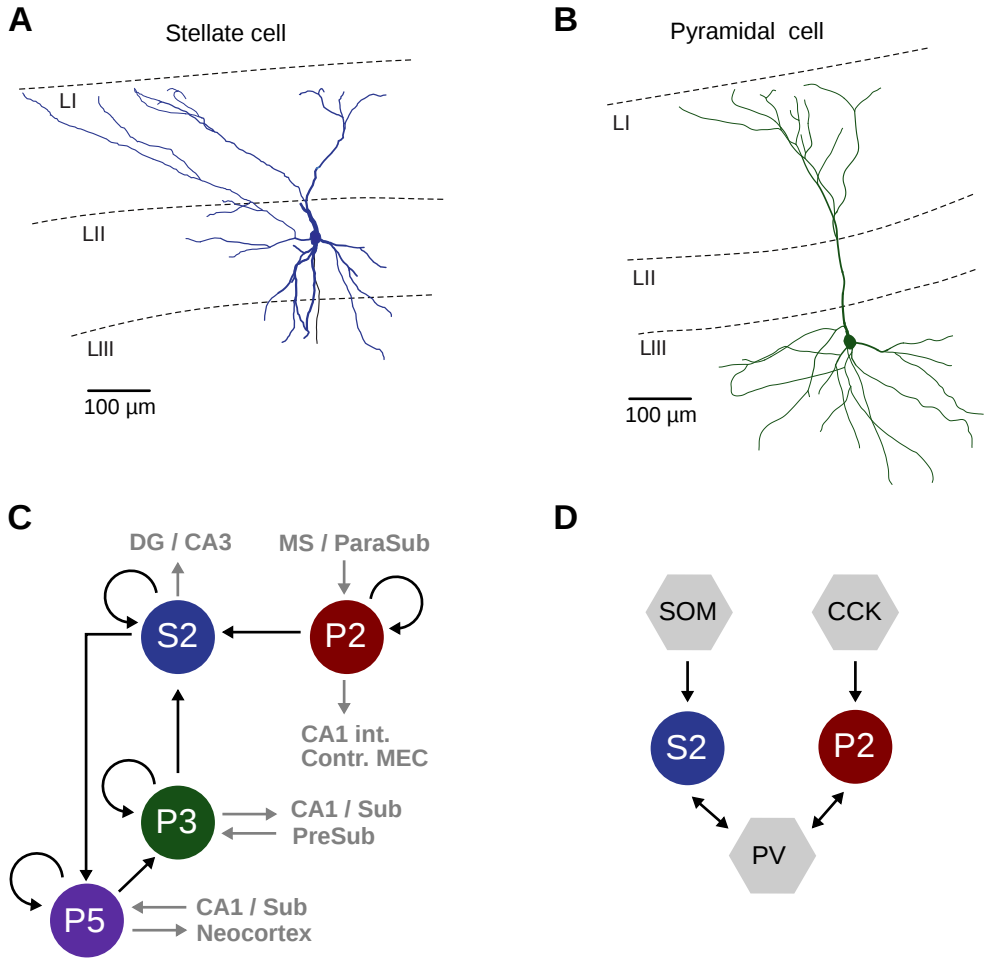

Figure 2.2. | **Microcircuits in the superficial layers of the MEC**. **A-B)** Histological reconstructions of a typical entorhinal stellate cell (A) and pyramidal cell (B). Dashed lines indicate boundaries between cortical layers. Image courtesy of Jochen Winterer. **C)** Diagram of local excitatory connectivity. S2, Layer II stellate cell; P2, Layer II pyramidal cell; P3, layer III pyramidal cell; P5, Layer V pyramidal cell. Black arrows: connections reported by Dhillon and Jones (2000); Van Haeften et al. (2003); Sürmeli et al. (2015); Winterer et al. (2017). Gray arrows: extrinsic projections (Ray et al., 2014; Tang et al., 2016, Figure 2.1C, D). **D)** Diagram of local inhibitory connectivity. CCK, cholecystokinin-positive interneuron; PV, parvalbumin-positive interneuron; SOM: somatostatin-positive interneuron. Diagram shows connections reported by Varga et al. (2010); Armstrong et al. (2016); Fuchs et al. (2016).

Cell-type specific connectivity within layer II is still debated: pair recordings *in vitro* revealed unidirectional projections from pyramidal to stellate cells (Winterer et al., 2017), although a developmental study suggests a projection in the opposite direction (Donato et al., 2017). Additionally, Winterer et al. (2017) found moderately large connectivity rates among layer II stellate cells (2.5%) and pyramidal cells (2.9%).

Interlaminar excitatory pathways have been also described in the MEC. Most of the entorhinal input to layer II arises from layer III (Beed et al., 2010), with layer III neurons targeting layer II stellate cells selectively (Winterer et al., 2017). Finally, layer V principal cells receive input from layer II (Sürmeli et al., 2015) and directly innervate layer III (Van Haeften et al., 2003; Kloosterman et al., 2003).

The projection patterns described above suggest that layer II stellate cells mediate connectivity pathways between the MEC and the hippocampal formation and process information across cortical layers (Figure 2.2C). By contrast, layer II pyramidal cells provide parasubicular inputs to the MEC and convey theta-rhythmic activity from the medial septum (Ray et al., 2014).

Inhibitory circuits

To understand how inhibition affects grid-cell activity (e.g., Buetfering et al., 2014; Miao et al., 2017), I shall now review the main types of inhibitory neurons in the MEC and their local connectivity.

The EC, similarly to other neocortical areas, contains three major classes of interneurons: parvalbumin (PV), somatostatin (SOM), and serotonin-3a-receptor (5HT3aR) expressing cells (Lee et al., 2010; Rudy et al., 2011). PV-positive interneurons—the most abundant interneuron type in the area—are typically fast-spiking basket or chandelier cells. Basket cells are characterized by a prominent axonal arborization that resembles a 'basket' surrounding the soma of the target cell, whereas chandelier cells are characterized by vertical clusters of synaptic boutons (candles) that contact the axon initial segment of the post-synaptic neuron. By targeting the perisomatic region directly, PV-positive inhibitory cells are in a unique position to control the action-potential output of their post-synaptic targets. In contrast, SOM-positive interneurons preferentially target the dendritic compartments of the post-synaptic cell, possibly gating it synaptic inputs. Finally, the third major interneuron class, the 5HT3aR expressing cells, show a diversity of morphological and electrophysiological profiles, and they are generally less-well characterized in comparison to PV or SOM cells (Vucurovic et al., 2010; Lee et al., 2010).

Local inhibitory circuits in the MEC have been best characterized within layer II (Figure 2.2D). PV-positive basket cells provide somatic inhibition to both stellate and pyramidal neurons (Beed et al., 2013; Couey et al., 2013; Armstrong et al., 2016; Fuchs

et al., 2016; Ray et al., 2017), whereas SOM-positive interneurons preferentially target stellate cells (Fuchs et al., 2016). On the other hand, CCK-positive interneurons (a subgroup of 5HT3AR expressing cells) inhibit pyramidal cells selectively (Varga et al., 2010; Armstrong et al., 2016; Fuchs et al., 2016).

2.2.2. Intrinsic properties of entorhinal neurons

A large number of experimental and theoretical studies, including the work of this thesis, postulate that grid-cell activity may arise via intrinsic cellular properties of entorhinal neurons (e.g., Burgess et al., 2007; Giocomo and Hasselmo, 2008a; Kropff and Treves, 2008; D'Albis and Kempter, 2017, Sections 4.2-4.3, and Chapter 5). To date, *in vitro* physiology work has largely focused on layer II stellate cells, because of their intrinsic resonant currents that have been linked to theta-rhythmicity and grid-cell firing (e.g., Burgess et al., 2007; Giocomo and Hasselmo, 2008a). In this section, I shall thus summarize the main intrinsic properties of entorhinal stellate cells, with comparisons to layer II pyramidal cells when data was available.

Sag and rebound potentials

A peculiar feature of entorhinal stellate cells is a pronounced membrane-potential sag in response to the injection of a subthreshold current (Alonso and Klink, 1993). That is, when a small current step is applied, the membrane potential reaches an early peak and then decays (or sags) to a lower steady-state level. Such sag responses have been linked to hyperpolarization-activated cation currents (I_h), which slowly activate upon hyperpolarization and slowly de-activate upon depolarization, opposing membrane-potential changes in both directions (Dickson et al., 2000). I_h currents also generate the so-called rebound after hyperpolarization potential, i.e., the membrane voltage overshoots the resting potential—possibly beyond spiking threshold—after a cell is released from prolonged hyperpolarization (Alonso and Klink, 1993). Sag and rebound responses are particularly prominent in stellate cells, but they are found in pyramidal cells too (Alonso and Klink, 1993). Sag responses are faster dorsally than ventrally (Giocomo et al., 2007), and they are considerably faster in stellate than in pyramidal cells (Winterer et al., 2017). I kindly refer the reader to (Hasselmo, 2014; Shay et al., 2016) for grid-cell models based on sag and rebound potentials.

Theta-frequency resonance

Possibly the most striking electrophysiological property of entorhinal stellate cells is the emergence of a rhythmic subthreshold oscillation following a constant depolarization of the membrane potential (Alonso and Llinás, 1989; Alonso and Klink, 1993). Remarkably,

the frequency of this oscillation matches the frequency of the hippocampal theta rhythm ($4 - 12$ Hz, Vanderwolf, 1969; Winson, 1974; Buzsáki, 2002), which has been implicated in spatial memory and grid-cell firing (Mitchell et al., 1982; Brandon et al., 2011; Koenig et al., 2011). Likewise, when a sinusoidally-modulated current is injected into a stellate cell, the sub-threshold membrane voltage resonates in the theta-frequency range (Haas and White, 2002; Erchova et al., 2004; Schreiber et al., 2004; Pastoll et al., 2012). Similarly to sag and rebound potentials, stellate-cell resonance properties depend on I_h currents (Robinson and Siegelbaum, 2003; Nolan et al., 2007; Haas et al., 2007), and they are modulated along the dorso-ventral axis of the MEC (with higher frequencies dorsally than ventrally; Giocomo et al., 2007).

In contrast to stellate cells, entorhinal pyramidal cells do not show intrinsic rhythmic oscillations *in vitro*, neither in the subthreshold nor in the suprathreshold regime (Alonso and Klink, 1993; Erchova et al., 2004; Schreiber et al., 2004). This supported the conjecture that stellate (but not pyramidal) cells may contribute theta-rhythmicity and grid-cell firing (Burgess et al., 2007; Giocomo and Hasselmo, 2008a). Yet recent experiments showed that spiking activity *in vivo* is more strongly theta-modulated in pyramidal than in stellate cells (Ray et al., 2014; Tang et al., 2014).

After-spike potentials

The action potential of a stellate cell is typically followed by an after-spike potential with three components: a shallow and fast after-spike hyperpolarizing potential (AHP, 2-5 ms), a brief after-spike depolarization, and a longer AHP (20-100 ms; Alonso and Klink, 1993; Pastoll et al., 2012). Similarly to other intrinsic properties, the time constants of these responses show a gradient along the dorso-ventral axis of the MEC, with faster responses dorsally than ventrally (Giocomo et al., 2007; Navratilova et al., 2012; Yoshida et al., 2013). Although stronger in stellate cells, after-spike potentials are found in pyramidal cells too (Alonso and Klink, 1993) as in other cortical projection neurons (Connors et al., 1982). I kindly refer the reader to (Navratilova et al., 2012) for a grid-cell model based on after-spike potentials.

Spike-rate adaptation

An important intrinsic property of entorhinal cells is spike-rate adaptation, i.e., when a depolarizing current is injected into a cell, its firing frequency rapidly decreases (or adapts) over time. In entorhinal stellate cells, the injection of a relatively large current step (e.g., 1 nA), causes the firing rate to drop logarithmically in the first tens of milliseconds and decay exponentially afterwards (Alonso and Klink, 1993). Typically, cells cease firing completely after 100-200 ms of constant current stimulation, but they may resume firing—forming peculiar spike clusters—if stimulated for longer times

(Alonso and Klink, 1993). Spike-rate adaptation is stronger dorsally than ventrally (Yoshida et al., 2013) and is more prominent in stellate than in pyramidal cells (Alonso and Klink, 1993). Spike-rate adaptation is particularly relevant for this thesis because it is part of the grid-cell model I propose in Chapter 5.

2.3. Chapter summary

In this chapter, I described the main components of the hippocampal system, a compound cortical region that is strongly implicated in spatial navigation and episodic memory. A key structure therein is the entorhinal cortex, which sits at the interface between the hippocampal formation and the neocortex. Multimodal sensory inputs reach the entorhinal cortex from widespread cortical areas mostly via the perirhinal and postrhinal cortices. Within the medial entorhinal cortex—where grid cells are most abundant—excitatory microcircuits with stereotyped projection patterns have been recently identified. Intrinsic cellular properties of entorhinal neurons, such as theta-frequency resonance and spike-rate adaptation, have been also characterized *in vitro*, suggesting possible mechanisms for the formation of grid-cell patterns. In the next chapter, I will introduce the reader to grid cells.

Chapter 3

Grid cells in the entorhinal cortex

Here, I review current empirical knowledge on grid cells. I describe the main properties of their firing patterns (Section 3.1), how they are affected by sensory inputs (Section 3.2), and how they relate to other spatial cells in the hippocampal system (Section 3.3).

3.1. Introduction

The entorhinal cortex has a long research history, mostly because well-studied brain disorders, such as epilepsy and Alzheimer's disease, often originate in this area (e.g., Van Hoesen et al., 1991; Spencer and Spencer, 1994). Yet nothing has boosted the interest in this brain region more than the discovery of grid cells.

The path that led to this influential breakthrough is marked the currently unsolved quest for the origin of place-cell activity in the hippocampus (see Section 2.1.2 and Moser et al. (2008) for an introduction to place cells). A key finding on this issue was that hippocampal place fields in CA1 were partially maintained after complete ablation of the CA3–CA1 connections, which suggested that the entorhinal cortex—the second major CA1 input—could provide spatially-tuned signals to the hippocampus (Brun et al., 2002). This idea was supported by earlier studies that reported a broad and irregular spatial firing in the MEC (Barnes et al., 1990; Quirk et al., 1992; Frank et al., 2000). Yet early entorhinal recordings mostly targeted the ventral part of the MEC, leaving the more dorsal areas largely unexplored. When the group of Edvard and May-Britt Moser recorded from the dorsal MEC, a truly unexpected discovery was made: not only spatial firing was much sharper than previously reported, but a single cell had multiple firing fields that were arranged in strikingly-regular triangular pattern in the environment (Fyhn et al., 2004; Hafting et al., 2005). These cells were termed grid cells by the experimenters (Figure 3.1A). Place and grid cells are currently one of the hottest topics in the field of systems neuroscience, and, for their discovery, John O'Keefe and Edvard and May-Britt Moser were awarded the 2014 Nobel Prize for

Physiology or Medicine.

3.1.1. Basic grid-cell properties

The spatial firing pattern of a grid cell is characterized by three fundamental properties: 1) the grid scale, i.e., the distance between two neighboring firing fields; 2) the grid orientation, i.e, the angle between one of the grid axis and a reference direction; and 3) the grid spatial phase, i.e., the two-dimensional spatial offset between the firing fields and a reference point (Figure 3.1B).

Nearby grid cells, i.e., cells simultaneously recorded from the same electrode, have similar scale and orientation, but scattered phases (Hafting et al., 2005). That is, grid scale and orientation are topographically organized in the cortex, whereas grid phase is largely non-topographic—similarly to the salt-and-pepper distribution of orientation preference in the rodent visual cortex (but see Heys et al., 2014, for a weak anatomical clustering of grid phases in linear tracks).

Grid scale and field size monotonically increase from the dorsal to the ventral MEC (Hafting et al., 2005, Figure 3.1C), mirroring a dorso-ventral gradient of many intrinsic properties of entorhinal cells (Section 2.2.2) and an increase of place-field size along the dorso-ventral axis of the hippocampus (Jung et al., 1994).

3.1.2. Modular organization of grid-cell activity

Only two years after grid-cell activity was first reported, Barry et al. (2007) added another interesting bit to this discovery. By sampling several dorso-ventral locations in the MEC, they observed that grid scale did not vary continuously in the tissue, but rather in discrete steps (Figure 3.1D). Interestingly, the ratio between two subsequent grid scales in a given animal was a constant number, about 1.7, which roughly matched theoretical predictions for the optimal coding of space with periodic signals (Stemmler et al., 2015; Wei et al., 2015). A similar scale ratio (in the range 1.4–1.7) was later reported by two other independent studies (Stensola et al., 2012; Krupic et al., 2015).

The finding by Barry et al. (2007) motivated the Moser's group to further investigate the anatomical organization of grid-cell properties within the cortex (Stensola et al., 2012). By recording hundreds of cells per animal across large portions of the MEC, Stensola et al. (2012) not only confirmed the step-like increase in grid scale observed by Barry et al. (2007), but also found that grid-cell activity was generally organized in a small number of discrete functional modules. Within a module, grid patterns had a similar scale, orientation, elliptic distortion, and theta-frequency modulation (Figure 3.1E), but the same properties varied substantially across modules. By sampling about a half of the dorso-ventral extent of the MEC, Stensola et al. (2012) counted four to five modules per animal, hinting that the total number of grid modules in the

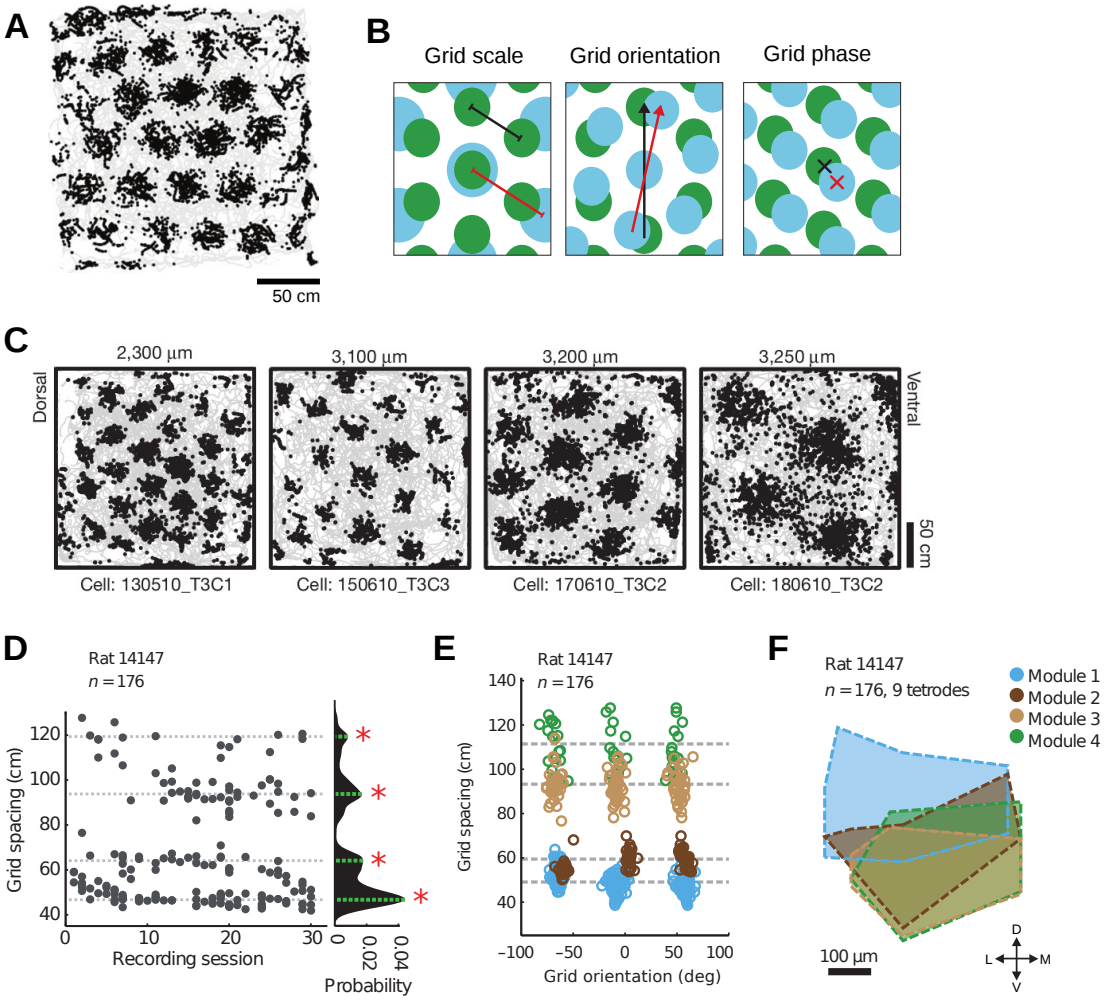

Figure 3.1. | Entorhinal grid-cell patterns: basic properties and modular organization. A) Spatial firing pattern of a grid cell recorded in the rat's MEC. The gray trace shows the trajectory of the animal foraging in a square enclosure. The black dots indicate the locations in which cell fired. **B)** Cartoons of two grid patterns (blue and green), illustrating the differences between grid scale (left), grid orientation (middle) and grid spatial phase (right). Grid scale, orientation, and phase for the green (blue) pattern are marked in black (red). **C)** Example grid firing patterns (black dots) recorded at successive dorsoventral locations in the MEC (left: dorsal, right: ventral). Dorsoventral location from the brain surface is indicated above the panels. Animal trajectory is depicted in gray. **D)** Left: Grid spacing for 176 grid cells recorded in a single animal. Right: kernel-density estimate of the distribution of grid spacings. Dashed lines and red stars indicate local peaks in the distribution (see also panels E and F). **E)** Scatterplots showing spacing and orientations of 176 grids recorded in the same animal (see panel D). Grid modules are color coded (see panel F). Orientation and spacing are measured for the three grid axes independently (three circles per grid). Gray dashed lines indicate the mean spacing of each module. **F)** Outlines of the anatomical extent of the grid modules in panels D and E depicted on a flat map of the medial half of the MEC. Images in panels A and B are adapted from (Moser et al., 2014) with permission from Nature Publishing Group. Images in panels D–F are adapted from (Stensola et al., 2012) with permission from Nature Publishing Group.

MEC is of the order of ten. Surprisingly, these modules also showed a large degree of overlap in anatomical space (Figure 3.1F), spanning multiple cortical layers and extending to pre- and parasubiculum (Stensola et al., 2012). Such a configuration suggest that functionally-identified grid-cell modules do not directly map onto anatomical cell clusters in the MEC, e.g., calbindin-positive (Fujimaru and Kosaka, 1996; Ray et al., 2014; Kitamura et al., 2014), zinc-positive (Slomianka and Geneser, 1997), or acetylcholinesterase-positive patches (Slomianka and Geneser, 1991); although more complex function-structure relationships may still exist (Naumann et al., 2018).

3.2. The role of sensory inputs in grid-cell firing

The discovery of grid cells raised a number of fundamental questions regarding the origin of spatial signals in the MEC. Do grid patterns emerge from a constellation of external sensory cues, or rather from a continuous integration of self-motion inputs? Are grids affected by the local geometry of the enclosure? Are they stable across environments? And do they provide a local or a global representation of space? In this section, I shall review the current knowledge on these topics as it emerged over more than ten years of investigation.

3.2.1. Self-motion inputs versus external sensory cues

When grid cells were first reported, Hafting et al. (2005) observed that grid fields remained stable in the dark, indicating that they could arise from the integration of idiothetic self-motion cues (e.g., vestibular signals or somatosensory feedback), which inform the animal about its own movements in the environment (see also Barry et al., 2012a). In the very same study, however, grid patterns were also found to rotate with polarizing visual cues in circular arenas, similarly to head-direction cells in presubiculum (Taube et al., 1990), and place cells in the hippocampus (Muller and Kubie, 1987). In the light of these findings, and inspired by previous models of place-cell activity (McNaughton et al., 1996; Samsonovich and McNaughton, 1997, Section 4.1.2), it was thus proposed that grid patterns may initially emerge from the integration of self-motion cues (path integration) and then get anchored to external landmarks with experience (McNaughton et al., 1991). It was hypothesized that grid cells could provide a "universal path integration-based neuronal map of the spatial environment" (McNaughton et al., 2006)—a conjecture that is currently dominating the field (Moser et al., 2017, see also Sections 4.1 and 4.2).

However, the relative contribution of self-motion inputs to grid-cell activity might have been initially overestimated: grid patterns in the dark could be sustained by olfactory cues or by somatosensory inputs resulting from the encounter of the arena

boundaries—factors that were initially not well controlled by experimenters (Hafting et al., 2005; Barry et al., 2012a). Consistent with this hypothesis is that in elevated arenas (Pérez-Escobar et al., 2016), and in conditions where olfactory cues are minimal (Chen et al., 2016), grid patterns are rapidly disrupted in darkness. Yet the relative role of idiothetic and allothetic signals may also differ across species: stable grids in the dark were observed in rats (Hafting et al., 2005; Barry et al., 2012a), but not in mice (Pérez-Escobar et al., 2016; Chen et al., 2016).

Carpenter et al. (2015) investigated how sensory inputs affect grid-cell activity in an interesting experiment. Rats were allowed to forage between two identical enclosures connected by a single hallway, similarly to a person visiting two identical hotel rooms on the same corridor. The authors hypothesized that, if grid-cell firing was dominated by self-motion cues, a global pattern (comprising the two rooms) should emerge. Alternatively, if external sensory inputs were prevailing, two local and identical patterns could emerge. The result was the following. Initially, identical patterns formed in the two rooms; however, after prolonged experience (2–3 weeks), the two grid maps merged, approaching a more global representation of space (Carpenter et al., 2015). This experiment suggests that external sensory inputs exert a strong control on grid-cell activity initially, and that self-motion cues may become important with experience—a conjecture that will be further elaborated in Chapter 6.

3.2.2. Influence of the local geometry of the enclosure

That grid-cell activity is controlled by external sensory cues is further supported by studies in which the local geometry of the environment is manipulated. Grid patterns are typically unaffected by the size or the aspect ratio of a familiar enclosure. However, when a square arena is suddenly rescaled along one or two directions, grid patterns either stretch accordingly (Barry et al., 2007), or rearrange their firing fields completely (Savelli et al., 2008). Similarly, grid fields are displaced locally when animals are allowed to move freely between previously discontinuous enclosures (Wernle et al., 2018). Remarkably, grid patterns deform coherently within (but not across) modules (Stensola et al., 2012; Yoon et al., 2013; Wernle et al., 2018), suggesting that functionally-independent grid-cell networks may coexist in the MEC.

The grid deformations reported by Barry et al. (2007) and Savelli et al. (2008) indicated that arena boundaries might control grid-cell activity more than previously expected. Indeed, Krupic et al. (2015) found that grids tend to align to the arena walls in square environments, and that this alignment is preserved after the arena is rotated within the room (despite prominent visual cues remaining stationary). Notably, grid orientations across rats seemed to cluster at a specific angular offset relative to the walls of the enclosure (Krupic et al., 2015; Stensola et al., 2015). The average angle to the walls

was 8.8 ± 4.7 degrees (mean $\pm$ s.d.) in the study by Krupic et al. (2015), and 7.2 ± 3.5 degrees in a second study by Stensola et al. (2015).

The observed grid alignment to the boundaries has also implications for the modular organization of grid-cell activity in the MEC: because grids of different modules oriented to the walls similarly, they could be either perfectly aligned to each other (0 degrees offset) or rotated by 30 degrees (i.e., 90 modulo 60 degrees). In fact, Krupic et al. (2015) observed that relative grid orientations across modules cluster around 0 and 30 degrees in single animals (with a large preference for 0 degrees), and that such cross-module alignment is maintained in non-square arenas, such as circles or hexagons (Krupic et al., 2015).

Finally, when grid cells were recorded in more complex arena shapes, such as the hairpin maze (Derdikman et al., 2009) or trapezoidal enclosures (Krupic et al., 2015, 2018), the distinctive triangular symmetry of grid-cell firing was largely lost—further supporting the idea that grid-cell activity is strongly influenced by environmental factors.

3.2.3. Grid cells across environments

Here, I describe how grid-cell patterns change across environments. To this end, I first explain the concept of 'remapping', a term originally introduced in the context of hippocampal place cells. Place-cell activity changes (or remaps) in two different ways depending on how the spatial context around the animal is manipulated. When a rat experiences two different boxes in the same room or similar boxes in two different rooms, hippocampal activity undergoes *global remapping*, i.e., place cells in one condition are either completely silent in the other condition or they have fields in different locations (Wills et al., 2005; Leutgeb et al., 2005, 2007). By contrast, when the physical recording room is kept constant but salient features of the arena are changed (such as wall colors, arena shape, or odors), place cells undergo *rate remapping*, i.e., firing fields are stable but firing rates vary substantially (Hayman et al., 2003; Leutgeb et al., 2005); see also (Latuske et al., 2017) for a recent review.

Fyhn et al. (2007) recorded the activity of entorhinal grid cells in conditions in which hippocampal firing undergoes either global or rate remapping. Interestingly, entorhinal activity was significantly more stable compared to the hippocampal counterpart, i.e., grid cells always maintained their spatially-periodic firing across contexts, even when place fields globally remapped. Additionally, when different boxes were tested in the same room, grid patterns shifted coherently within modules, i.e., cells recorded from the same electrode had similar phase-shift offsets. Similarly, when two identical boxes were tested in different rooms, grids rotated and translated, but, within modules, their relative-phase relationships remained constant—a phenomenon termed 'coherent

remapping' in the literature (Fyhn et al., 2007; Yoon et al., 2013). Furthermore, in conditions where place fields underwent rate remapping, grids either remained completely stable (Fyhn et al., 2007) or shifted coherently (Marozzi et al., 2015), although peak firing rates within fields varied substantially (Diehl et al., 2017; Ismakov et al., 2017).

Place- and grid-cell remapping are typically studied in environments that are already familiar to the animal. In this case, grid fields maintain similar size and spacing across conditions (Fyhn et al., 2007; Marozzi et al., 2015; Diehl et al., 2017; Ismakov et al., 2017). By contrast, grid patterns expand and become less regular in novel environments, i.e., in arenas that are experienced by the animal for the first time (Barry et al., 2012a).

3.3. Relation between grid cells and other spatial cells

Grid cells are only one of many spatially-modulated cell types in the hippocampal system, including place cells (O'Keefe, 1976), head-direction cells (Rank, 1984), border cells (Solstad et al., 2008), speed cells (Kropff et al., 2015), goal-direction cells (Sarel et al., 2017), and object-vector cells (Hoydal et al., 2018). In this section, I shall summarize the key findings that relate grid-cell activity to the best-characterized spatial neurons in the hippocampal system, i.e., place, head-direction, and border cells (Figure 3.2; see Hartley et al., 2014, for a review).

3.3.1. Relation between grid and place cells

The link between grid- and place-cell activity is rather intricate, and it is currently unsettled whether grid cells contribute to place-cell firing or vice versa (see, e.g., Bush et al., 2014, for a review). The hippocampus is directly innervated by the superficial layers of the MEC (Section 2.1.1, Figure 2.1), where grid cells are most abundant (Hafting et al., 2005; Sargolini et al., 2006). This led to the hypothesis that hippocampal place fields are generated from entorhinal grid fields (e.g., Solstad et al., 2006; Rolls et al., 2006; Cheng and Frank, 2011). Consistently, it was found that hippocampal neurons receive monosynaptic connections from entorhinal grid cells (Zhang et al., 2013), and that genetic manipulations of entorhinal neurons affect the spatial scale of both grid and place fields (Mallory et al., 2018). Yet two lines of evidence indicate that grids are not required for place fields to emerge. First, place cells develop prior to grid cells during ontogenesis (Langston et al., 2010; Wills et al., 2010). Second, place fields are largely spared by entorhinal lesions (Brun et al., 2008; Van Cauter et al., 2008; Schlesiger et al., 2018) and by disruption of grid-cell activity via septal inactivation (Koenig et al., 2011); although place fields could still emerge from degraded grids (Azizi et al., 2014).

On the other hand, grid fields could be generated (Section 4.3) or anchored to the

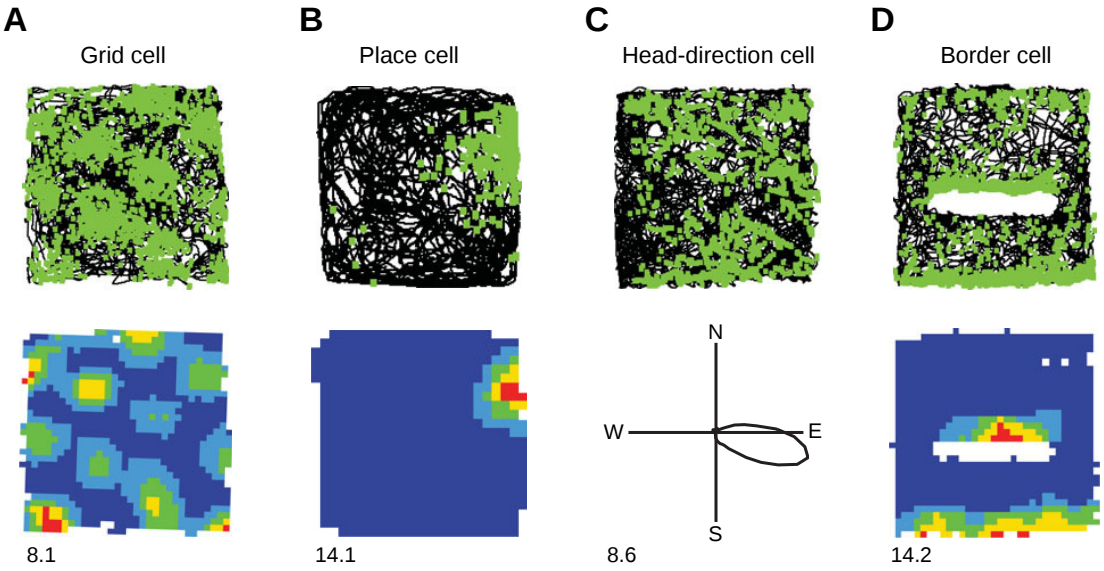

Figure 3.2. | Spatial cells in the hippocampal system. Examples of four spatially-modulated cell types in the hippocampal system: grid cell (A), place cell (B), head-direction cell (C), and border cell (D). At the top row, the locations at which spikes occurred (green squares) are superimposed on the movement trajectory of the animal (black line). At the bottom row, average firing-rate maps (A,B,D) or average directional tuning curves (C) are shown. In the firing-rate maps, warm colors depict high rates. Peak firing rates in spikes/s are indicated at the bottom-left corner. For the experiment in panel D, a bar parallel to the south wall was inserted in the arena to demonstrate the emergence of a second field on the north-side of the barrier. Grid, head-direction, and border cells were recorded in a 1 x 1 square box with 50 cm high walls. The place cell in panel B was recorded in a 62 x 62 cm arena. N, north; E, east; S, south; W, west. Figure adapted from (Hartley et al., 2014) with permission from the Royal Society.

environment (Section 4.1.5) via hippocampal place fields. In fact, the deep entorhinal layers (layers V and VI) receive direct synaptic projections from the hippocampus (Tamamaki and Nojyo, 1995; Sürmeli et al., 2015, Section 2.1.1), and hippocampal inactivations (Bonnevie et al., 2013) or hippocampal lesions (Fyhn et al., 2004) disrupt grid-cell firing.

3.3.2. Relation between grid and head-direction cells

Head-direction (HD) cells are neurons that fire according to the orientation of the animal's head in the environment, regardless of the animal's physical location (Figure 3.2C; Taube et al., 1990). Originally discovered in dorsal presubiculum (Rank, 1984; Taube et al., 1990), HD cells have been later reported in multiple brain areas, including the anterior thalamus (Taube, 1995a), the parasubiculum (Taube, 1995b), and the entorhinal cortex (Sargolini et al., 2006). Within the MEC, HD cells are most abundant in layers III, V, and VI, where grid cells are also found (Sargolini et al., 2006); and neurons with

conjunctive grid-by-HD tuning have also been recorded in the same layers (Sargolini et al., 2006). Because HD cells could provide self-motion information to grid cells, their discovery has been particularly relevant for computational models of grid-cell activity (Section 4.1). Additionally, the existence of conjunctive grid-by-HD cells has been predicted by theorists (McNaughton et al., 2006).

Interestingly, inactivations or lesions of the anterior thalamic nucleus (where HD cells are abundant), degrade both HD fields and grid-firing patterns in the MEC (Winter et al., 2015). On the other hand, septal inactivations—which reduce theta rythmicity throughout the hippocampal system—disrupt grid patterns but spare the HD signal, i.e., grid patterns disappear, conjunctive grid-by-HD cells become pure HD cells, and HD cells are largely unaffected (Brandon et al., 2011; Koenig et al., 2011).

3.3.3. Relation between grid and border cells

Border cells are neurons that selectively fire as the animal approaches the boundaries of an enclosure (Figure 3.2D). They are found in the MEC (Solstad et al., 2008; Savelli et al., 2008) and in parasubiculum (Boccara et al., 2010), and they are functionally related to the boundary-vector cells in subiculum, i.e., neurons that fire whenever the animal is at a certain distance to a boundary (Barry et al., 2006; Lever et al., 2009).

Because arena boundaries influence the symmetry and alignment of grid fields (Section 3.2), it is likely that border cells provide input to grid cells. Accordingly, border cells have been proposed to correct path-integration errors in models of grid-cell activity (Hardcastle et al., 2015, Section 4.1.5).

3.4. Chapter summary

In this chapter, I introduced the reader to grid cells. First, I described the basic geometrical properties of grid patterns and their modular organization in the cortical tissue. I then discussed the relative role of self-motion and sensory inputs in controlling grid-cell firing. I argued that grid-cell activity is primarily driven by environmental rather than self-motion cues, and that the latter could only be used for short time stretches. Finally, I discussed how grid cells relate to other spatially-modulated neurons in the hippocampal system. In the next chapter, I will present the main theories for the origin of grid-cell activity in the medial entorhinal cortex.

Chapter 4

Theories of grid-pattern formation

Here, I discuss the most-important computational models for the origin of grid-cell patterns. I shall focus on three classes of models: continuous-attractor models, which emphasize the role of recurrent connectivity (Section 4.1); oscillatory-interference models, which emphasize the role of rhythmic oscillations (Section 4.2); and single-cell plasticity models, which emphasize the role of external sensory cues and learning (Section 4.3).

4.1. Continuous-attractor models

Continuous-attractor networks (CANs) have been very influential in the field of computational neuroscience and they underlie one of the most-prominent theories for the origin of grid-cell patterns. The core idea behind those models is that a properly-wired recurrent network can sustain a continuum of stable states in the activity of its neurons (Amari, 1977; Taylor, 1999). Because all states are neutrally stable, small external perturbations can push the network to a nearby configuration so that the network activity at any time reflects the cumulative effects of the external inputs onto the system (Wu et al., 2008).

As CANs effectively function as neural integrators (Robinson, 1989), they have been implicated in models of many brain structures that track continuously-changing stimuli. For example, CAN models have been used to track eye position in the oculomotor system (Cannon et al., 1983; Seung, 1998), movement direction in the motor cortex (Georgopoulos et al., 1993), stimulus orientation in the visual cortex (Ben-Yishai et al., 1995), memory cues in prefrontal cortex (Camperi and Wang, 1998), head direction in the thalamus (Skaggs et al., 1995; Blair, 1996; Zhang, 1996), and animal position in the hippocampal system (McNaughton et al., 1996; Samsonovich and McNaughton, 1997; McNaughton et al., 2006; Fuhs and Touretzky, 2006).

Attractor states can be represented by CANs in two different ways: with either a 'rate code' or a 'location code' (Goldman et al., 2010). In a rate-coding scenario, the network

maintains a state of graded persistent firing, and the stimulus value is encoded by the firing rate of the neurons—as for example in the oculomotor system (Cannon et al., 1983; Seung, 1998). By contrast, in a location-coding scenario, the network generates an activity packet or 'bump' at the population level, and the stimulus value is encoded by the bump location on the neural sheet, i.e., by the identity of the active neurons. CANs with a location code have been used in models of head-direction, place, and grid cells, and they are thus particularly relevant for the aim this thesis. Because all CANs with a location code share the same basic architecture, I shall first introduce the network setup for the simplest scenario, the head-direction system.

4.1.1. Continuous-attractor models of head-direction cells

Figure 4.1A schematically illustrates the architecture of a typical CAN model for head-direction cells (e.g., Skaggs et al., 1995; Blair, 1996; Zhang, 1996; Stringer et al., 2002b). Neurons that will acquire head-direction tuning are arranged on a ring (panel A1). Nearby neurons on the ring are strongly excitatorily connected, whereas distant neurons are weakly excitatorily connected or inhibited. Such a connectivity pattern generates a localized activity bump on the ring (panel A1 warm ball colors). The bump is then moved in concert with the rotation of the animal's head by two groups of 'rotation neurons': clockwise-rotation (inner ring in panel A2) and anticlockwise-rotation neurons (inner ring in panel A3). Rotation neurons are conjunctively activated by head-direction cells (black arrows) and by vestibular inputs (dashed arrows). The former inform about the head's current direction, the latter about the head's turning velocity, i.e., when the animal turns its head right (left), clockwise (anticlockwise) rotation neurons are activated, and the activation strength is proportional to the turning speed. Crucially, rotation neurons project back to the head-direction ring asymmetrically, e.g., a clockwise-rotation cell that is innervated by head-direction cell ϕ projects back to head-direction cell $\phi + \Delta$, where Δ is a small angle in the clockwise direction.

Such a network effectively generates a head-direction signal by integrating angular velocity from the vestibular system. At any given time, head orientation can be decoded from the location of the activity bump on the (outer) ring, where each cell acquires a bell-shaped directional tuning curve. Importantly, the topology of the network matches the nature of the signal that is encoded, i.e., cells are connected on a ring because head-direction is a one-dimensional periodic variable. Note, however, that a ring-like connectivity does not necessary imply a topographical arrangement of preferred head orientations on the neural tissue—which in fact is not observed in mammals (Taube et al., 1990; Taube, 1995a). CAN models of head-direction cells, first theorized in the 1990s, have recently found direct experimental support from imaging studies in the fly brain (Seelig and Jayaraman, 2015; Kim et al., 2017; Turner-Evans et al., 2017).

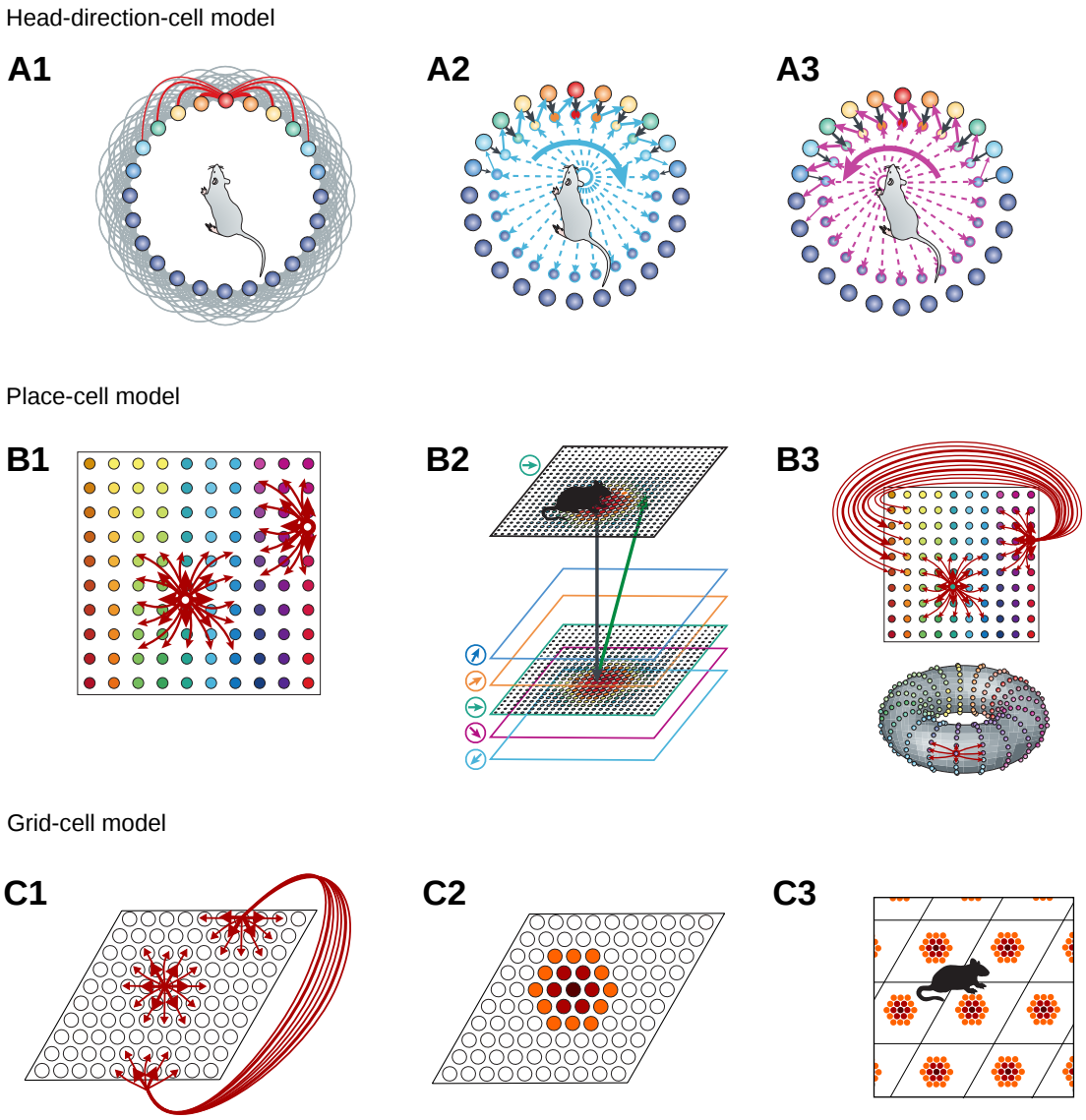

Figure 4.1. | Continuous attractor models of head-direction, place, and grid cells. **A)** Head-direction-cell CAN model. A1) Head-direction cells (balls) are topologically arranged on a ring. Nearby neurons are excitatorily connected (red and gray lines), and an activity bump emerges on the ring (warm ball colors). Bump rotation is achieved by means of two additional layers of neurons: clockwise-rotation neurons (inner ring in A2, blue outlines) and anticlockwise-rotation neurons (inner ring in A3, pink outlines). Blue (pink) dashed arrows indicate vestibular inputs for clockwise (anticlockwise) rotation. See main text for details. **B)** Place-cell CAN model. Two-dimensional extension of the ring attractor in A. Place cells (balls) are arranged on a plane with connection strengths declining with distance (B1). An activity bump emerges on the plane (B2). Conjunctive place-by-velocity cells (B2, colored layers) project to the place-cell plane asymmetrically, and shift the bump in concert with the animal movements. To avoid edge effects, connections with periodic boundaries may be considered (B3 top), which result in a connectivity with toroidal topology (B3 bottom). **C)** Grid-cell CAN model. The network setup is analogous to the place-cell model in B with periodic boundaries (B3), but the neuronal plane is tilted, i.e., with axes that are 60 degrees apart (C1). Conjunctive grid-by-velocity cells (not shown) shift the population activity bump (C2) in concert with the animal movements, and a triangular pattern emerges in physical space (C3). Images in panels A and B adapted from (McNaughton et al., 2006) with permission from Nature Publishing Group.

4.1.2. Continuous-attractor models of place cells

Place-cell activity in the hippocampus has been also explained by CAN models (Figure 4.1B, McNaughton et al., 1996; Samsonovich and McNaughton, 1997; Stringer et al., 2002a; Conklin and Eliasmith, 2005; McNaughton et al., 2006). Similarly to head-direction models, feedback connections within a network sustain an activity bump at the population level (Figure 4.1B1), and conjunctive place-by-velocity cells shift the bump on the neural sheet according to the animal's movements in physical space (Figure 4.1B2). As a result, a spatially-localized packet of neural activity (a place field) also emerges in single cells. Because spatial location on a plane is intrinsically two-dimensional, the network connectivity is now defined on a two-dimensional neuronal sheet. To update the bump position on the sheet, at least three layers of conjunctive place-by-velocity cells are required, i.e., a minimum of three movement directions shall be represented independently. Yet the model would work equally well with an arbitrary number of conjunctive layers, given that head direction—a proxy for movement direction—seems to be smoothly represented in the brain (Taube et al., 1990).

A critical point of this theory is that the size of the neural sheet should somehow scale with the size of the environment, which seems biologically unrealistic. To circumvent this problem, Samsonovich and McNaughton (1997) used periodic boundary conditions on the neural sheet, i.e., they assumed a recurrent connectivity with toroidal topology (Figure 4.1B3). This solution, however, directly implies that hippocampal place fields would periodically re-occur in large environments, which has never been observed experimentally (Fenton et al., 2008).

4.1.3. Continuous-attractor models of grid cells

Since periodic tuning has been discovered in the entorhinal cortex (Hafting et al., 2005), CAN models of hippocampal place cells have been quickly adapted to explain the activity of grid cells (McNaughton et al., 2006; Fuhs and Touretzky, 2006). One of the problems that remained to be solved concerned the geometry of the spatial patterns: the toroidal connectivity proposed by Samsonovich and McNaughton (1997) generated fields on a square grid of locations, whereas grid fields are typically arranged on a triangular lattice (Figure 3.1A). To overcome this issue, one could change the connectivity structure of the network to a twisted-torus topology (Guanella et al., 2007), where the 'twist' consists in tilting the axes of the neuronal sheet from 90 degrees (Figure 4.1B3) to 60 degrees (Figure 4.1C1). Such a tilt, affects the way the population-activity bump moves on the periodic network, and generates a triangular grid at the single-cell level (Figure 4.1C2, C3). The grid scale (Section 3.1.1) could be controlled by the speed-modulation gain of the conjunctive cells, i.e., the cells that shift the bump on the neuronal sheet with the animal's movements. For high speed gains, the bump

moves fast across the network, and small grid scales arise. Conversely, for lower speed gains, the bump moves slower, and larger grids emerge. Equivalently, keeping the speed gain constant, periodic networks of different sizes could also produce multiple grid scales.

An alternative way of implementing a CAN model for grid cells is to generate multiple bumps of activity at the population level (Fuhs and Touretzky, 2006; Burak and Fiete, 2009; Couey et al., 2013; Widloski and Fiete, 2014). If such population bumps are already arranged on a triangular lattice, rigidly shifting the bumps with the animal's movements also produces triangular grids at the single-cell level. But how to generate multiple activity bumps in the first place? The answer lies in the connectivity profile of the network. If the recurrent-connectivity strength smoothly decays with distance on the neural sheet (e.g., following a Gaussian profile), only a single bump of activity can emerge at the population level—as in all examples of Figure 4.1. However, if the connectivity function follows a Mexican-hat (short-range excitation and long-range inhibition) or a Lincoln-hat (a step-like change in connection strength) profile, the network can sustain multiple activity bumps simultaneously. Mathematically, it is the Fourier spectrum of the connectivity function that matters: if the spectrum peaks at zero frequency (e.g., for a Gaussian or a cosine function) a single bump will emerge; but if it peaks at a positive frequency (e.g., for a Mexican hat or Lincoln hat function), multiple bumps can arise. Furthermore, when a non-linearity is present in the system, such bumps can also spontaneously arrange on a triangular grid on the neuronal sheet—similarly to the spontaneous pattering of chemical concentrations in reaction-diffusion systems (Turing, 1952; Borckmans et al., 2002). Multi-bump attractor networks have the advantage that they can form single-cell triangular grids even in the absence of periodic connectivity. However, due to edge effects, such aperiodic networks are inherently poor velocity integrators, and they are unlikely to generate stable grids at the single-cell level (Burak and Fiete, 2009, see Section 4.1.5).

4.1.4. The problem of learning continuous attractors

CAN models of grid-cell activity crucially rely on specific connectivity patterns between the neurons. But how such a connectivity structure could arise in the biological system? It is often assumed that Mexican-hat connectivity (i.e., short-range excitation and long-range inhibition) reflects physical distance between neurons. This assumption, however, predicts that grid spatial phases shall be topographically organized in the cortex, which is not observed experimentally (Hafting et al., 2005).

To overcome this problem, McNaughton et al. (2006) proposed a two-step learning process that guides the formation of CANs during development. The idea is the following. Initially, a population-activity pattern emerges in a 'teaching layer' of neurons,

which is transient and will disappear in the adult animal. Such an activity pattern arises via a Mexican-hat connectivity that mirrors physical distance between the neurons. Synaptic noise then randomly translates the pattern across the teaching layer, and provides downstream neurons (or modules) with the right correlation structure to form periodic recurrent connections locally (McNaughton et al., 2006). Crucially, these modules do not require a topographic organization of grid phases in the tissue, which is consistent with experimental data (Hafting et al., 2005). The developmental model by McNaughton et al. (2006) is theoretically intriguing, but has been challenged mathematically (Burak and Fiete, 2006, 2009) and has found little experimental support so far. I kindly refer the reader to (Fuhs and Touretzky, 2006) for a similar developmental model based on spontaneous cortical waves.

Widloski and Fiete (2014) proposed an alternative method to learn grid-cell CANs. In contrast to previous proposals (McNaughton et al., 2006; Fuhs and Touretzky, 2006), their model requires spatial exploration during ontogenesis. As the animal explores a familiar space, an initially random network is activated by both place-selective and self-motion inputs. With these inputs, temporally-asymmetric synaptic plasticity generates a network connectivity similar to the one of hard-wired CANs (see, e.g., Figure 4.1B2). After all connections have been learned, synaptic plasticity is turned off and the dynamics of the mature network is probed with self-motion inputs alone. In this 'activation' phase, the recurrent connections sustain multiple bumps of activity that shift coherently with the animal's movements in physical space.

Here, I would like to point out a few issues of this model. First, as highlighted by the authors, the population-activity pattern moves across the mature network inconsistently with the spatially-tuned inputs that were active during learning (Widloski and Fiete, 2014). This inconsistency is advertised as a feature of the model, because it makes the functioning of the mature network independent of the initial learning experience. It seems however conceptually odd that feed-forward and recurrent inputs shall be in contradiction in a cue-rich and familiar environment where learning initially occurred. Second, because the learned connectivity structure is aperiodic, the model is inherently brittle to any source of noise present in the system (Burak and Fiete, 2006, 2009). Finally, because pattern formation depends on the details of the synaptic plasticity rules, it remains unclear whether the proposed mechanism is realistic in biological terms. I kindly refer the reader to Chapter 6 for a more thorough discussion of this model.

4.1.5. Integration accuracy and anchoring

To obtain a single-cell grid from a population pattern, CAN models require precise integration of the animal's velocity and heading. Yet any integrating system is inherently prone to errors if left to run freely without external calibrating signals. In the case of

grid-cell CANs, an integration error of the order of the grid spacing (tens of centimeters) is sufficient to disrupt single-cell patterns completely. Burak and Fiete (2009) estimated that a stochastic periodic network of $\sim 10^4$ neurons accumulates errors at a rate of ~ 1 cm/min—suggesting that velocity integration shall be corrected (or reset) every few minutes of exploration. In the biological system, however, errors may accumulate faster, due to smaller network sizes, inhomogeneities in the connectivity patterns, or spatially-irregular feed-forward inputs. Additionally, because velocity integration is a relative process, coherent grid firing across exploration sessions requires anchoring the network activity to stable landmarks in the environment.

Landmark-based error correction and anchoring have been often discussed only in abstract terms in the grid-cell literature (e.g., Welinder et al., 2008) with just a few studies attempting to model such processes explicitly (Guanella et al., 2007; Pastoll et al., 2013; Hardcastle et al., 2015). A possibility is that grid-cell CANs are constantly calibrated by hippocampal place cells (Guanella et al., 2007; Pastoll et al., 2013)—similarly to visual cues calibrating CAN models of head-direction cells (McNaughton et al., 1991; Skaggs et al., 1995; Zhang, 1996). As the animal explores a new environment, Hebbian synaptic plasticity could rapidly form associations between simultaneously active place and grid cells. In this case, place-cell firing could correct CAN integration errors within sessions, and anchor single-cell grids across sessions (Guanella et al., 2007; Pastoll et al., 2013). However, because place-cell activity remaps across environments (Section 3.2), this solution requires a 'one-shot learning' of many place-to-grid associations in each environment.

A more parsimonious solution to the same problem uses border-cell inputs for calibration (Hardcastle et al., 2015). In fact, Hardcastle et al. (2015) have empirically observed that the spatial dispersion of grid-cell spikes is consistent with integration drifts predicted by CAN models, and that such drifts are reduced after the encounter of an arena boundary (Hardcastle et al., 2015). Additionally, grid-cell firing appears more precise in the direction perpendicular to the latest-encountered boundary, consistent with the hypothesis that border cells could provide error-correction signals to grid cells (Hardcastle et al., 2015). Because border cells are typically stable across spatial contexts (Solstad et al., 2008), border-based calibration does not need rapid learning in novel environments. Border inputs, however, provide spatial information only in the direction perpendicular to the boundary, which is not sufficient to anchor single-cell grids across sessions.

4.1.6. Evidence in favor and against grid-cell attractor models

A fundamental feature of all attractor models of grid-cell activity is that spatial firing is driven by self-motion cues. This is in line with the facts that rodents can use self-motion

cues to navigate for short distances (Etienne and Jeffery, 2004; Gil et al., 2018; Tennant et al., 2018); that grid fields can in some cases persist in darkness (Hafting et al., 2005; Barry et al., 2012a; Chen et al., 2016; Pérez-Escobar et al., 2016, Section 3.2.1); that head-direction and conjunctive cells have been observed in the MEC (Sargolini et al., 2006; Kropff et al., 2015); and that grid patterns are degraded after suppressing head-direction signals in the thalamus (Winter et al., 2015). Finally, the attractor theory is consistent with the modular organization of grid-cell activity and its coherent remapping across environments (Stensola et al., 2012; Yoon et al., 2013, and Section 3.1.2); though the same effects could arise from a recurrent coupling that develops after the initial pattern-formation process (see Chapter 6).

On the other hand, attractor models require a sophisticated neural architecture whose developmental origin is unclear (Section 4.1.4); they need anchoring to the physical space (Section 4.1.5); and they cannot explain how grid fields are distorted by external sensory cues and the geometry of the environment (Sections 3.2.1 and 3.2.2).

4.2. Oscillatory-interference models

A second class of grid-cell models is based on temporally-interfering oscillations in single neurons (e.g., O'Keefe and Burgess, 2005; Burgess et al., 2007; Burgess, 2008; Blair et al., 2008; Zilli and Hasselmo, 2010; Navratilova et al., 2012).

4.2.1. The dual-oscillator theory of place-cell activity

Oscillatory-interference (OI) models of grid-cell activity originate from a theory proposed in the early 1990s to explain the origin of place fields and their relationship with the hippocampal theta rhythm. At that time, O'Keefe and Recce (1993) observed that place cells preferentially activate in bursts, with an inter-burst interval that is slightly shorter than a period of the extracellular theta. When the animal enters a place field, spikes typically occur at the trough of the theta rhythm, and they gradually advance in phase as the animal runs through the field—a phenomenon termed 'phase precession' by the authors (Figure 4.2A, black trace and bottom plot); see also (Skaggs and McNaughton, 1996; Huxter et al., 2008; Schmidt et al., 2009).

O'Keefe and Recce (1993) proposed that place fields and phase precession could be simultaneously generated by the interference of two oscillations impinging onto a single cell, i.e., a cell-intrinsic oscillation (green trace in Figure 4.2A) and a global theta-frequency oscillation (red trace in Figure 4.2A). If the cell-intrinsic frequency was slightly higher than the theta frequency, a beating interference pattern could emerge in the membrane potential of the cell. The pattern's envelope would set the place-field temporal extent, while its carrier frequency would provide the right spike timing to

generate phase precession (Figure 4.2A, blue and pink traces). This temporal beating pattern would then be translated into a spatial signal by modulating one of the two oscillations with the animal's running speed (O'Keefe and Recce, 1993; Lengyel et al., 2003).

The 'dual-oscillator' theory by O'Keefe and Recce (1993) may thus explain the emergence of place fields and phase precession at the same time. However, as highlighted in Figure 4.2A, such a mechanism also generates fields that periodically repeat in the environment. To solve this problem, O'Keefe and Recce (1993) assumed that constructive interference (beating) only occurs within restricted spatial regions (the place fields), whereas destructive interference (equal frequencies but opposite phases) occurs in the rest of the environment. Yet it remained unclear how such a frequency switch (from constructive to destructive interference) could be implemented in the brain. This issue was easily circumvented with the discovery of periodic firing in the entorhinal cortex.

4.2.2. Oscillatory-interference models of grid-cell firing

Because the dual-oscillator model intrinsically generates periodic patterns, it was immediately proposed as a mechanism for grid-cell activity (O'Keefe and Burgess, 2005). Consistently, it was soon observed that grid cells—like place cells—also show phase precession, both in linear tracks (Hafting et al., 2008; Reifenstein et al., 2012) and in open fields (Climer et al., 2013; Jeewajee et al., 2014; Newman and Hasselmo, 2014; Reifenstein et al., 2014).

The original dual-oscillator model was then extended to explain grid-cell activity in two-dimensional environments (Burgess et al., 2007). It was initially assumed that, within a single cell, a somatic oscillation at theta frequency interfered with multiple dendritic oscillations at slightly higher frequencies. To obtain a planar interference pattern, the frequencies of the dendritic oscillators were modulated by the running speed and direction of the animal (velocity-controlled oscillators, VCO). It was soon realized, however, that dendritic and somatic oscillations inevitably lead to phase locking in single cells (Remme et al., 2010). Therefore, VCOs were later modeled as separate groups of afferent units, often termed 'theta cells' in the literature (Burgess, 2008; Blair et al., 2008).

To generate a triangular grid, VCOs must be tuned to animal's movement directions differing by multiples of 60 degrees (Figure 4.2B). Two VCOs are sufficient to generate grid-like patterns (Figure 4.2B), but six are required to account for omnidirectional phase precession as observed experimentally (Figure 4.2C; Climer et al., 2013; Jeewajee et al., 2014; Reifenstein et al., 2014). Yet none of the proposed configurations can fully account for the path-dependent features of phase precession that were empirically described (Reifenstein et al., 2014).

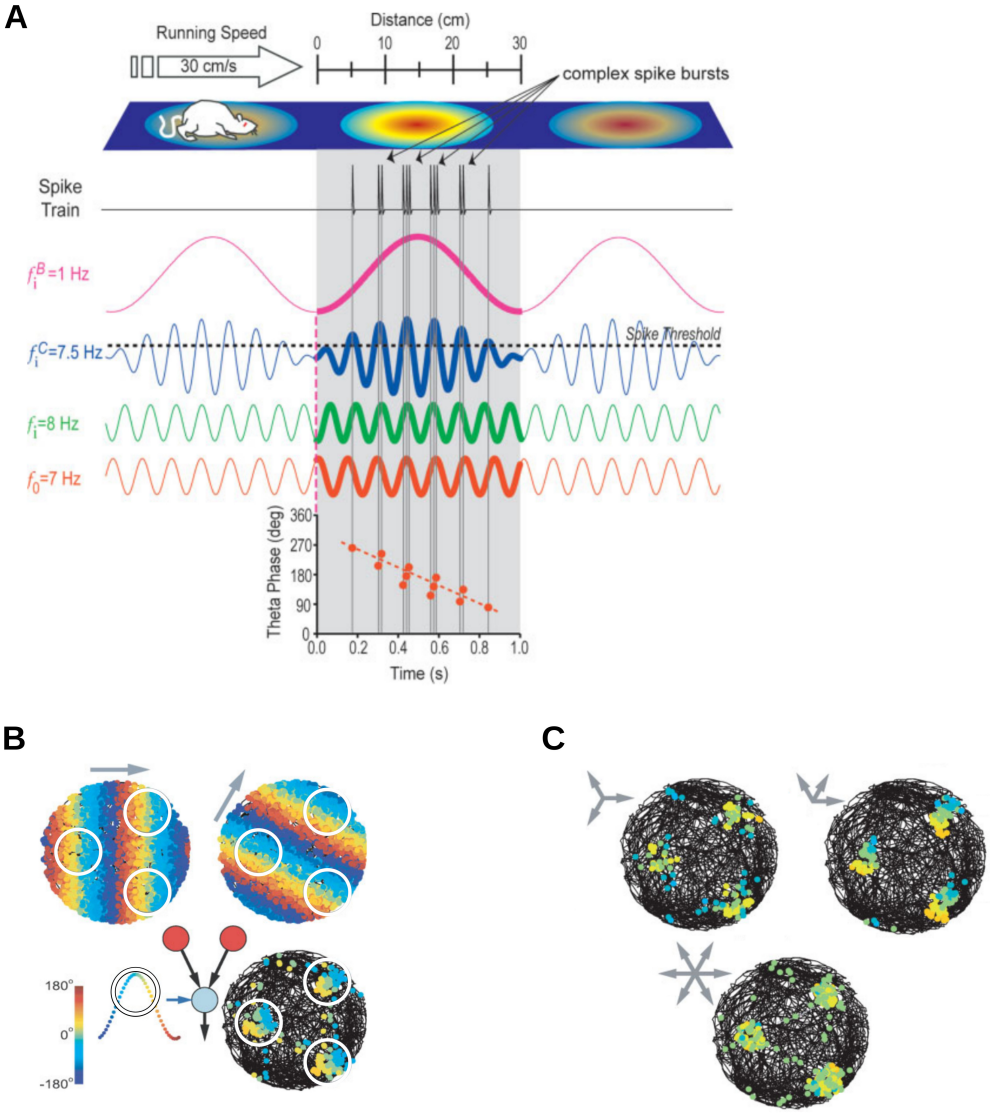

Figure 4.2. | **Oscillatory-interference models for place and grid cells**. **A)** As a virtual rat runs across a linear track at constant speed, two oscillators with frequencies $f_0 = 7$ Hz (red) and $f_i = 8$ Hz (green) interfere with each other generating a 1 Hz beating pattern riding on a 7.5 Hz carrier (blue). Place fields are defined by the envelope of the oscillation (magenta). The cell fires when its membrane potential (blue) exceeds a spike threshold (dashed line). Action potentials are generated in bursts (black lines) at the 7.5 Hz carrier frequency and undergo phase precession against the baseline 7 Hz theta rhythm (bottom). **B)** Two velocity-controlled oscillators (VCOs, red circles) and a baseline theta rhythm (blue arrow) impinge onto a grid cell (light blue circle). The top plots show the spike phases of the two VCOs with respect to the baseline theta in a circular environment (see colorbar). The VCOs have different preferred running directions (gray arrows). The grid cell fires whenever it receives coincident input from the two VCOs (i.e., similar spike phases, locations denoted by white circles). The bottom-right plot shows the output spike locations and phases over the virtual-rat trajectory (black). Note that grid-cell spike phases are location dependent, i.e., the cell does not show omnidirectional phase precession. **C)** Spike locations and phases of a grid cell in a circular arena for three different configurations of VCOs (gray arrows). Omnidirectional phase precession is obtained only with six VCOs that are 60 degrees apart. Image in panel A is reproduced from (Blair, 1996; Blair et al., 2008) with permission from John Wiley and Sons. Images in panel B and C are adapted from (Burgess, 2008) with permission from John Wiley and Sons.

A weakness of OI models is that the characteristic 60-degree periodicity of grid-cell patterns must be manually inserted into the system via the preferred VCOs directions. Additionally, when more than two VCOs are employed—as needed to explain phase precession—the phases of such oscillators must precisely match to form triangular patterns. It has been proposed that VCOs with a 60-degree angular offset could spontaneously emerge via self-organizing process (Burgess et al., 2007; Mhatre et al., 2012), but such a scenario has never been explicitly simulated in OI models.

Another problem of the OI theory is that grid-like periodicity could be rapidly disrupted by noise, e.g., due to the incoherence of the theta rhythm beyond a few oscillation cycles (Welinder et al., 2008; Zilli et al., 2009). Although network synchronization mechanisms could potentially alleviate the problem (Zilli and Hasselmo, 2010), it remains unclear whether such mechanisms could be still effective in biologically realistic networks.

4.2.3. Evidence in favor and against grid-cell interference models

OI models are consistent with entorhinal neurons exhibiting sub-threshold oscillations *in vitro* (Alonso and Llinás, 1989; Alonso and Klink, 1993); with grid-cell activity being theta-modulated and phase-precessing *in vivo* (Hafting et al., 2005; Domnisoru et al., 2013; Schmidt-Hieber and Häusser, 2013); and with velocity-modulated theta cells (reminiscent of VCOs) being experimentally identified (Welday et al., 2011). Additionally, the intrinsic resonance of stellate cells exhibits a dorso-ventral frequency gradient that mirrors an increase of the grid spacing along the same anatomical axis (Giocomo et al., 2007). Finally, abolishing theta modulation originating in the medial septum disrupts grid-cell firing in the MEC (Brandon et al., 2011; Koenig et al., 2011).

Nevertheless, experimental evidence against OI models also exists: intrinsic resonance is not required for grid patterns to emerge (Giocomo et al., 2011); grid-like activity is observed without theta modulation in crawling bats (Yartsev et al., 2011); and the membrane-potential dynamics of grid cells is inconsistent with a purely interference-based mechanism (Domnisoru et al., 2013; Schmidt-Hieber and Häusser, 2013). The latter is probably the strongest piece of evidence against OI models, as will be further discussed in the next section.

4.2.4. Hybrid interference-attractor models

OI models posit that grids emerge from the temporal interference between multiple VCOs and a baseline theta rhythm, generating a beating signal in the membrane potential of the output cell. This predicts that grid-cell firing is mainly driven by the amplitude modulation of the intracellular theta (Figure 4.3A). By contrast, CAN models generate a bump of neural activity that is displaced across the network by

velocity-modulated signals (conjunctive cells). This predicts that output spikes are generated by periodically occurring ramps of synaptic depolarization (Figure 4.3B).

To test these predictions, Domnisoru et al. (2013) and Schmidt-Hieber and Häusser (2013) recorded the membrane voltage of entorhinal grid cells during virtual spatial exploration. A typical intracellular trace is shown in Figure 4.3C. Within firing fields, a clear depolarizing ramp was present in the membrane potential of the cell (red trace). Yet both the membrane voltage and the output spikes were also clearly theta modulated, and the intracellular power typically increased within firing fields (green trace). Although, these features are partially consistent with both OI and CAN models, further analysis revealed that the depolarizing ramp (and not the theta power) is highly predictive of the output spiking of the cell (Domnisoru et al., 2013; Schmidt-Hieber and Häusser, 2013), which favored CAN over OI models. It also became clear, however, that the intracellular theta-modulation observed experimentally was not explained by CAN models, leading to consider hybrid interference-attractor models of grid-cell firing (Schmidt-Hieber and Häusser, 2013; Bush and Burgess, 2014).

The main idea of hybrid interference-attractor models is to synaptically connect an ensemble of grid cells whose spatial firing is driven by temporal interference. The assumed connectivity profile is symmetric, i.e., synaptic connections are strongest between cells with similar spatial phases and gradually decrease in strength for cells with distant phases. Asymmetric connections with conjunctive cells (as assumed by CAN models) are instead not required, because the shift of the activity bump is directly controlled by the VCOs projecting to the recurrent network. Therefore, spatial periodicity and intracellular theta modulation are generated by the interference mechanism, whereas the sub-threshold depolarization ramps are generated by the recurrent connections. This mechanism predicts that removing lateral connectivity between grid cells shall abolish the depolarization ramps but spare a sub-threshold beating pattern in the membrane voltage (Figure 4.3A).

Because hybrid models are driven by temporal interference they also generate phase precession (Schmidt-Hieber and Häusser, 2013; Bush and Burgess, 2014). However, intracellular theta modulation and phase precession could be also generated by CAN models alone. In fact, Navratilova et al. (2012) proposed that an attractor network could generate grid fields and phase precession by using only theta-modulated conjunctive cells and intrinsic synaptic currents (Navratilova et al., 2012). To reproduce realistic grid scales, however, the model required very long synaptic delays (300–600 ms) and intrinsic currents whose temporal dynamics was modulated by the animal's running speed.

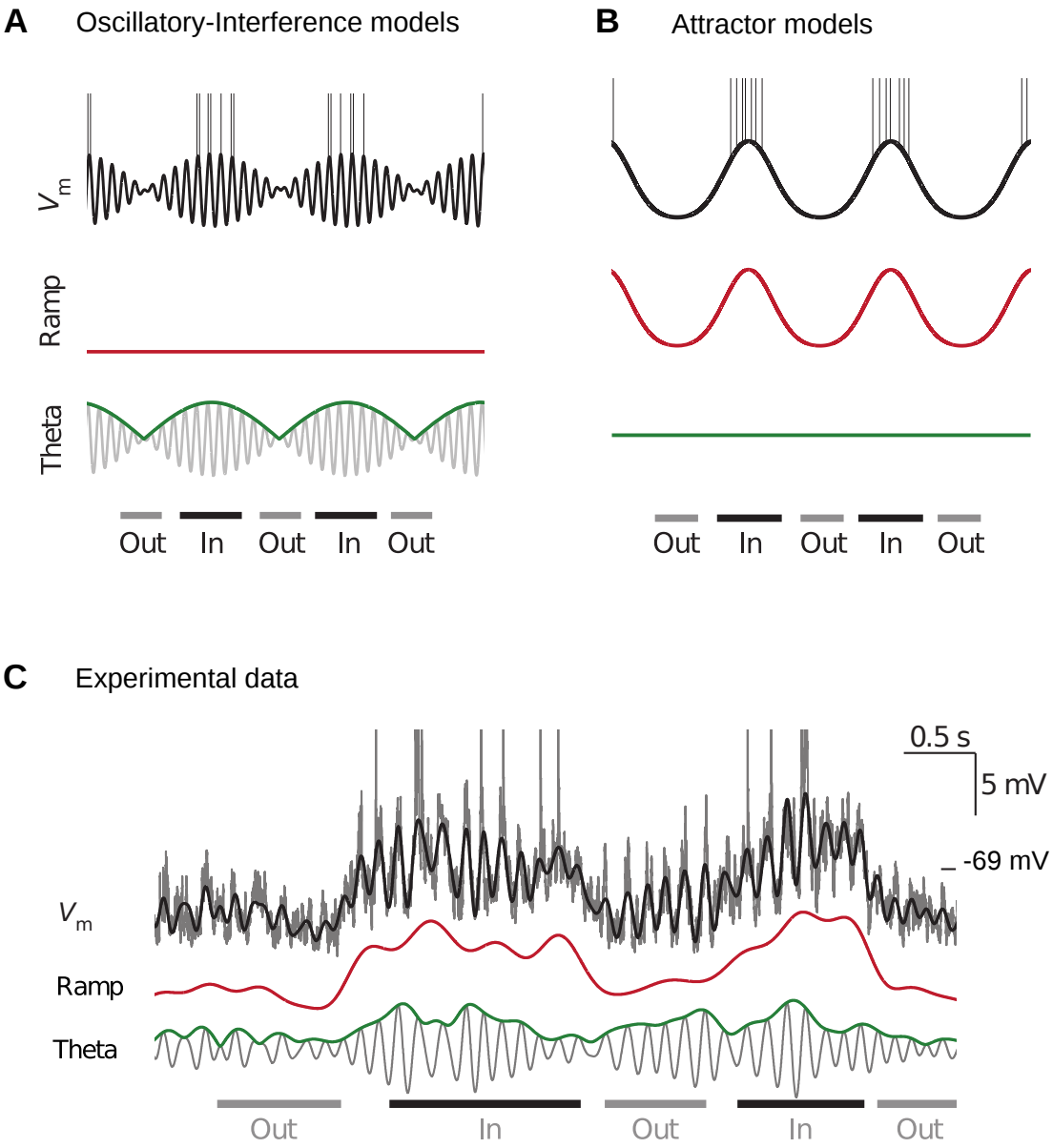

Figure 4.3. | **Membrane potential dynamics of grid cells**. **A)** Schematics of a grid-cell membrane potential as predicted by oscillatory-interference (OI) models. The membrane potential (V_m, black trace) is decomposed into a low-frequency ramp component (red) and a theta-frequency component (gray, envelope in green). Dashes at the bottom indicate periods of in-field activity (black) and out-of-field activity (gray). **B)** Schematics of a grid-cell membrane potential as predicted by continuous attractor network (CAN) models. Colors as in A. **C)** The intracellular membrane potential (gray, spikes truncated) of a typical entorhinal grid cell recorded on a virtual linear track (Domnisoru et al., 2013). The signal is decomposed in a theta-frequency component (gray) and a low-frequency ramp (red, see also panels A, B). The sum of the two components (black overlay) closely approximates the original signal. Images in panels A-C adapted from (Domnisoru et al., 2013) with permission from Nature Publishing Group.

4.3. Single-cell plasticity models

Finally, I discuss a third class of models in which grid patterns originate at the single-cell level via spatially-tuned feed-forward inputs and Hebbian synaptic plasticity (Kropff and Treves, 2008; Mhatre et al., 2012; Castro and Aguiar, 2014; Stepanyuk, 2015; Dordek et al., 2016; D'Albis and Kempter, 2017; Monsalve-Mercado and Leibold, 2017; Weber and Sprekeler, 2018). The model I propose in Chapter 5 also belongs to this category. Differently to the theories discussed earlier (Sections 4.1 and 4.2), periodic spatial patterns emerge here from the integration of external sensory inputs rather than self-motion signals.

4.3.1. The Kropff-and-Treves model

Kropff and Treves (2008) suggested that grid patterns may originate when synaptic excitation and cell-intrinsic mechanisms compete to control the output firing of a neuron. In particular, a competition could arise between spatially-selective synaptic inputs and spike-rate adaptation—a cell-intrinsic process that hinders a neuron to fire for long time stretches. In this case, Hebbian synaptic plasticity at the input synapses could imprint a spatially-periodic pattern in the output activity of the cell.

The model architecture is depicted in Figure 4.4A. A single cell (the output neuron, black disk) receives spatially-selective input from an ensemble of presynaptic cells (white disks). For simplicity, each input neuron is active in a single location of the environment, i.e., the input activity resembles that of hippocampal place cells. Input synapses are plastic and the output neuron is subject to spike-rate adaptation. Figure 4.4B illustrates the pattern-formation mechanism. As the animal traverses a place field (dark-red disk), the corresponding presynaptic input is active, the output firing rate increases, and the related synaptic weight is potentiated by Hebbian plasticity. After a short time, however, the output firing rate decreases due to spike-rate adaptation, i.e., the neuron is less active in a ring of positions around the current location of the animal (light-blue disk). Inputs with place fields within this 'adaptation ring' are uncorrelated with the output spiking and the corresponding weights are depressed. As the output activity recovers, a new input place-field association can be learned (light-red disk). Such a learning process reaches an equilibrium when the output place fields are roughly equally-distributed in the arena. Assuming isotropic spatial exploration (i.e., roughly circular adaptation rings), an optimal equilibrium in terms of coverage is obtained when the output fields are arranged in a lattice of equilateral triangles (Figure 4.4C; Thue, 1910)—a configuration that is reminiscent of grid-cell patterns (Hafting et al., 2005). Figure 4.4D shows the results of such a learning process in a simulation by Kropff and Treves (2008).

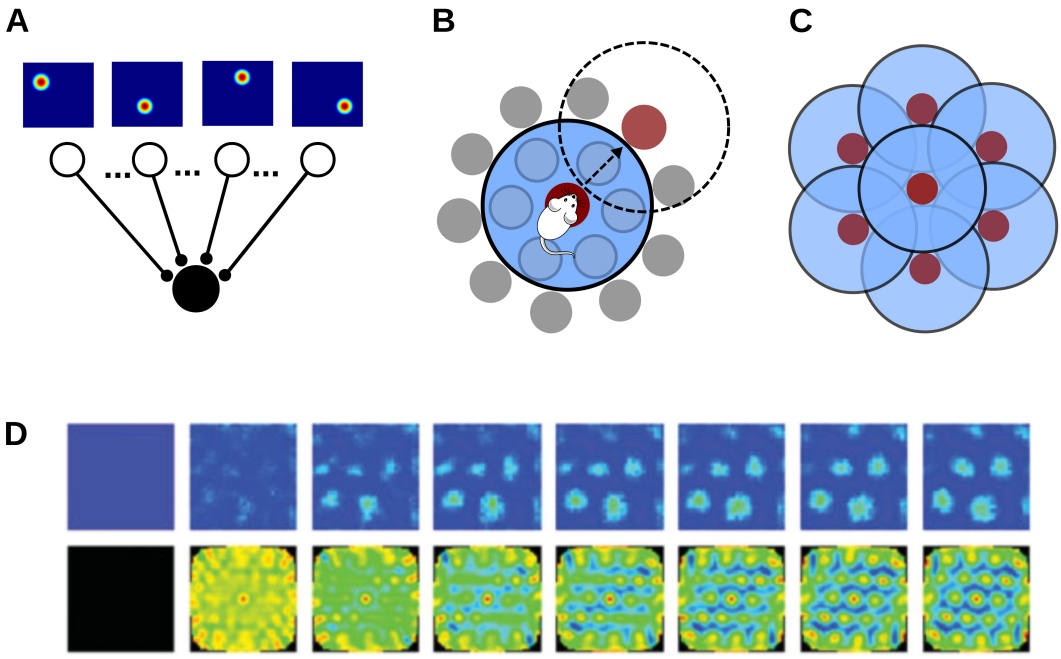

Figure 4.4. | **The Kropff-and-Treves model of grid-cell firing**. **A)** Schematics of the model architecture. A single cell (black disk) receives synaptic input from spatially-tuned neurons (white disk). Input spatial-tuning maps are shown at the top (warm colors denote high activity). **B)** Illustration of the pattern-formation mechanism. Input place fields cover the entire environment evenly (small disks). The dark-red disk at the center denotes a place field at the current location of the animal. Initially, input and output firing rates are high and the associated synaptic weight is potentiated by the Hebbian rule. Subsequently, the output neuron decreases firing due to spike-rate adaptation, i.e., the neuron is less active within an 'adaptation ring' around the current location of the animal (light-blue disk). After recovery from adaptation, the animal has moved outside the adaptation ring, and a new association is learned (light-red disk). The dashed circle denotes the following adaptation ring. **C)** Illustration of a possible equilibrium of the learning process. The hexagonal packing is a theoretical optimum. **D)** Example of the evolution of the synaptic weights during learning. Top row: each map shows the synaptic weights to the output neuron at a given time point (eight snapshots). Weights are arranged according to the positions of the corresponding input place-field centers (warm colors denote high weights). Bottom row: spatial auto-correlograms of the maps at the top row. Image in panels D reproduced from (Kropff and Treves, 2008) with permission from Wiley and Sons.

The spacing between the potentiated input fields (hence the scale of the output pattern) is controlled by the time constant of adaptation and the average running speed of the animal. Because learning requires a considerable amount of time to stabilize (hours of spatial exploration), variations of the animal's running speed on short time scales (e.g., seconds) do not hinder the pattern-formation process (Kropff and Treves, 2008). Additionally, the assumption of a place-like input tuning can be relaxed to inputs with multiple firing fields randomly distributed in the environment (Kropff and Treves, 2008).

Grid alignment through recurrent collaterals

Leveraging on an inherently single-cell mechanism, the Kropff-and-Treves model in its simplest form cannot explain the common alignment of grid patterns within modules (Hafting et al., 2005; Fyhn et al., 2007; Stensola et al., 2012, Section 3.1.2). To overcome this problem, recurrent collaterals and head-direction tuning have been later added to the model (Kropff and Treves, 2008; Si et al., 2012). Assuming that cells initially show head-direction tuning, recurrent collaterals are set up such that cells with similar directional selectivity are strongly interconnected. This recurrent connectivity guides the learning of the feed-forward connections such that grids with similar orientations emerge at the output (Kropff and Treves, 2008; Si et al., 2012). Note, however, that because cells maintain their directional selectivity after learning, this model is more suitable to describe the origin of conjunctive grid-by-head-direction cells rather than the one of pure grid cells.

The common grid orientation at the output is random in circularly-symmetric environments, but may be biased towards the edges in quadrangular arenas (Si et al., 2012)—which is consistent with empirical observations (Krupic et al., 2015; Stensola et al., 2015). Si et al. (2012) proposed that such an alignment to the boundaries could stem from anisotropies of the animal's running speed and direction during development. Additionally, Si et al. (2012) showed that the same recurrent collaterals that align the grids could also constrain the learning of the feed-forward inputs across environments, thus explaining the phenomenon of coherent remapping observed experimentally (Fyhn et al., 2007). Finally, Si and Treves (2013) proposed that a recurrent connectivity suitable for grid alignment could be learned in an unsupervised manner.

In summary, these studies predict that head-direction tuning is required for grid alignment, and that this alignment may arise at a later stage during development, i.e., after the initial formation of the grid patterns.

4.3.2. Other single-cell plasticity models

Following the initial proposal by Kropff and Treves (2008), several other investigators considered models of the same type (e.g., Mhatre et al., 2012; Castro and Aguiar, 2014; Stepanyuk, 2015; Dordek et al., 2016; Monsalve-Mercado and Leibold, 2017; Weber and Sprekeler, 2018). In all these models, grid-spatial tuning originates from a single-cell process based on spatially-selective inputs and Hebbian synaptic plasticity. However, the mechanisms underlying the emergence of spatial periodicity and triangular symmetry differ across models and often generate distinct predictions. In the following, I shall briefly summarize the core ideas underlying those models and highlight their most distinctive features.

A model based on spatially-periodic stripe cells

Mhatre et al. (2012) proposed that grid patterns could arise by combining inputs that activate in spatially-periodic bands in the environment. Such inputs, termed 'stripe cells' by the authors, resemble a subset of spatially-tuned cells empirically observed in the MEC and parasubiculum (Krupic et al., 2012), but see also (Navratilova et al., 2015) for a possible confound. Starting from a population of stripe cells with mixed orientations and phases, a Hebbian plasticity rule selects inputs that form a triangular grid at the output (Mhatre et al., 2012). In fact, activity bands with a 60-degree angular offset tend to co-occur more frequently than bands at any other orientation, as the authors demonstrate geometrically in their study (Mhatre et al., 2012). Therefore, the model explains how triangular symmetry could self-organize provided that spatial periodicity is already given at the input. But how to obtain spatially-periodic stripes in the first place? Mhatre et al. (2012) hypothesized that stripe cells are generated by one-dimensional ring attractors, each of which is tuned to a different movement direction of the animal (see also Section 4.1.1). This requires however a large number of cells and precise synaptic connectivity. Extensions of this model have been studied by Pilly and Grossberg (2012, 2013).

Models based on the dynamics of the synaptic-plasticity rule

Castro and Aguiar (2014) and Stepanyuk (2015) studied grid-cell models in which spatial periodicity stems from specific features of the synaptic plasticity rules considered. In particular, it was assumed that synaptic plasticity could switch between long-term potentiation (LTP) and long-term depression (LTD) based on the level of presynaptic firing, alike to the BCM theory (Bienenstock et al., 1982). Such a rule, combined with non-linear homeostatic terms (Stepanyuk, 2015), or inhibitory activity (Castro and Aguiar, 2014), could then generate grid-like patterns from place-selective inputs. Here,

differently from the Kropff-and-Treves model, intrinsic adaptation is not required for learning. Yet the hypothesized plasticity mechanism is rather complex, and it has not been described in the entorhinal cortex. The models by Castro and Aguiar (2014) and Stepanyuk (2015) predict that the grid scale shall be affected by interfering with the dynamics of synaptic plasticity.

A model based on a balance between excitation and inhibition

An alternative explanation for the origin of grid-cell activity is provided by the work of Weber and Sprekeler (2018). In their model, both excitatory and inhibitory inputs impinge onto a single cell. Both input types are spatially-selective and the spatial tuning of inhibition is broader than the one of excitation, e.g., excitatory input place fields are larger than the inhibitory ones. All synaptic connections are plastic: the excitatory weights undergo classical Hebbian learning, whereas the inhibitory weights are subject to an homeostatic learning rule that aims at maintaining the output firing-rate constant over time. However, because inhibition is more broadly tuned than excitation, a perfect balance between the two inputs cannot be obtained, and periodic spatial patterns emerge at the output. This model is conceptually similar to the Kropff-and-Treves model, i.e., spatial periodicity stems from a competition between persistent excitation and the reluctance of a neuron to fire continuously in the environment. In the model by Weber and Sprekeler (2018), however, the neuron stops firing due to synaptic inhibition rather than spike-rate adaptation. This model predicts that inhibitory plasticity is required for grid-cell learning, that inhibition is more broadly tuned than excitation, and that the tuning-width difference between excitation and inhibition scales with the grid spacing in the MEC.

A model based on phase precessing inputs

Monsalve-Mercado and Leibold (2017) proposed a grid-cell model based on place-cell activity and theta-phase precession (see Section 4.2.1). Similarly to other models in this section, a population of place cells project to a single neuron, and the input synapses are plastic. In this case, however, the input place cells also exhibit theta-phase precession, i.e., they activate in sequences within single cycles of the theta oscillation. For this setup, Monsalve-Mercado and Leibold (2017) showed that grid-like patterns could emerge at output. The model is supported by numerical simulations and rigorous analytical calculations, but a clear intuition of how phase precession generates periodic firing is not provided. The model predicts that abolishing phase precession in hippocampal place cells shall disrupt grid-cell firing in the MEC.

A model based on principal-component analysis

Finally, I discuss the work of Dordek et al. (2016). Compared to other models presented previously, this study attempts to explain the origin of grid-cell patterns at an algorithmic level, without making specific assumptions on the underlying cellular mechanisms. The core insight is that periodic triangular patterns could arise by performing a principal-component analysis (PCA, Hotelling, 1933) of bell-shaped inputs. Because it has been previously shown that PCA can be implemented by Hebbian plasticity in single neurons (Oja, 1982), Dordek et al. (2016) proposed that the same principle could underlie grid-pattern formation in the MEC. Importantly, to generate a periodic output, it is assumed that the input spatial tuning curves are 'zero mean' and resemble Mexican-hat functions. Yet because firing rates are non-negative, this assumption is biologically implausible. The authors discuss that effectively zero-mean inputs could be obtained by either temporal filtering or by spatially-tuned inhibition. In the former case, the model is similar to the proposal by Kropff and Treves (2008), in the latter, it is similar to the proposal by Weber and Sprekeler (2018). In either case, the study by Dordek et al. (2016) does not seem to generate any additional experimental prediction.

4.4. Chapter summary

In this chapter, I presented the main computational models explaining the emergence of grid-cell patterns. I grouped the proposed theories in three main classes: continuous-attractor models, oscillatory-interference models, and single-cell plasticity models. Attractor models posit that grids emerge from recurrent-network dynamics and self-motion inputs. By contrast, oscillatory-interference models propose that grids arise via theta oscillations that produce beating patterns in the intracellular voltage of a neuron. Finally, single-cell plasticity models posit that spatially-selective inputs and Hebbian learning generate spatially-periodic tuning at the single-cell level. In this thesis, I study a grid-cell model which belongs to this third category (Chapter 5). My own research work aims at increasing the biological realism of these models while giving deeper analytical insights on spatial pattern formation. In the next chapter, I describe my model in details.

Part II.

Original research

Chapter 5

The origin of grid-cell patterns

Here, I describe a novel computational model for the origin of grid-cell patterns. The contents of this chapter have been published in a peer-reviewed journal article entitled "A single-cell spiking model for the origin of grid-cell patterns" (D'Albis and Kempter, 2017).

5.1. Introduction

Grid cells are neurons of the medial entorhinal cortex (MEC) tuned to the position of the animal in the environment (Hafting et al., 2005; Rowland et al., 2016). Unlike place cells, which typically fire in a single spatial location (O'Keefe, 1976; Moser et al., 2008), grid cells have multiple receptive fields that form a regular triangular pattern in space. Since their discovery, grid cells have been the object of a great number of experimental and theoretical studies, as they are thought to support high-level cognitive functions such as self-location (e.g., Fiete et al., 2008a; Mathis et al., 2012a), spatial navigation (e.g., Bush et al., 2015; Stemmler et al., 2015; Horner et al., 2016), and spatial memory (Ólafsdóttir et al., 2016; O'Neill et al., 2017). Nevertheless, to date, the mechanisms underlying the formation of grid spatial patterns are yet to be understood (Chapter 4).

The attractor-network theory proposes that grid fields could arise from a path-integrating process, where bumps of neural activity are displaced across a low-dimensional continuous attractor by self-motion cues (e.g., McNaughton et al., 2006; Fuhs and Touretzky, 2006; Guanella et al., 2007; Burak and Fiete, 2009, Section 4.1). The idea that self-motion inputs could drive spatial firing is motivated by the fact that mammals can use path integration for navigation (Etienne and Jeffery, 2004), that speed and head-direction signals have been recorded within the MEC (Sargolini et al., 2006; Kropff et al., 2015), and that grid firing fields tend to persist in darkness (Hafting et al., 2005; Barry et al., 2012a); but see (Chen et al., 2016; Pérez-Escobar et al., 2016) and Section 3.2. However, grid-cell activity may rely also on non-visual sensory inputs—such as olfactory or tactile cues—even in complete darkness (Maaswinkel and

Whishaw, 1999). Additionally, the attractor theory alone cannot explain how grid fields are anchored to the physical space (Section 4.1.5), and how the properties of the grid patterns relate to the geometry of the enclosure (Barry et al., 2007; Savelli et al., 2008; Krupic et al., 2016, 2018; Wernle et al., 2018, and Sections 3.2.2 and 4.1.6).

A different explanation for the formation of grid-cell activity is given by the so-called oscillatory-interference models (e.g., O'Keefe and Burgess, 2005; Burgess et al., 2007, Section 4.2). In those models, periodic spatial patterns are generated by the interference between multiple oscillators whose frequencies are controlled by the velocity of the animal. Speed-modulated rhythmic activity is indeed prominent throughout the hippocampal formation in rodents and primates (Vanderwolf, 1969; Winson, 1974; McFarland et al., 1975; Sławińska and Kasicki, 1998), particularly within the theta frequency band (4-12 Hz). Additionally, reduced theta rhythmicity disrupts grid-cell firing (Brandon et al., 2011; Koenig et al., 2011), and grid-cell phase precession (Hafting et al., 2008) is intrinsically generated by interference models; but see (Reifenstein et al., 2014). Despite their theoretical appeal, however, these models cannot explain grid-cell activity in the absence of continuous theta oscillations in the bat (Yartsev et al., 2011), and they are inconsistent with the grid-cell membrane-potential dynamics as measured intracellularly (Domnisoru et al., 2013; Schmidt-Hieber and Häusser, 2013, Section 4.2.3).

Here, I focus on the idea that grid-cell activity does not originate from self-motion cues, but rather from a learning process driven by external sensory inputs. In particular, it was proposed that grid patterns could arise from a competition between persistent excitation by spatially-selective inputs and the reluctance of a neuron to fire for long stretches of time (Kropff and Treves, 2008, Section 4.3.1). In this case, Hebbian plasticity at the input synapses could imprint a periodic pattern in the output activity of a single neuron. Spatially-selective inputs, i.e., inputs with significant spatial information, are indeed abundant within the MEC (Tang et al., 2014; Sun et al., 2015; Diehl et al., 2017) and its afferent structures (Cacucci et al., 2004; Hargreaves et al., 2005, 2007; Boccara et al., 2010; Tang et al., 2016). And spike-rate adaptation, which is ubiquitous in the brain (La Camera et al., 2006), could hinder neuronal firing in response to persistent excitation.

Kropff and Treves (2008) explored this hypothesis by means of a computational model. The emergence of grid-like patterns was demonstrated with theoretical arguments and with numerical simulations of a rate-based network. However, because of a relatively abstract level of description, the outcomes of the model could not be easily confronted with experimental data. Specifically, the simulations included a network-level normalization mechanism that constrained the mean and the sparseness of the output activity, and it remained unsettled whether grid patterns could emerge in a single-cell scenario. Additionally, the synaptic weights did not obey Dale's law. And

the robustness of the model was not tested against shot noise due to stochastic spiking. Finally, the link between the numerical simulations and the underlying mathematical theory remained rather loose.

To overcome these issues, I propose here a single-cell spiking model based on similar principles as the model by Kropff and Treves (2008), but that is, on the one hand, more biologically realistic, and on the other hand, better suited for mathematical treatment. Importantly, I show that grid patterns can emerge from a single-cell feed-forward mechanism needless of any network-level interaction. To increase biological plausibility, I consider a stochastic spiking neuron model, and I constrain the synaptic weights to non-negative values (Dale's law). Finally, by studying the model analytically, I quantitatively predict the requirements for grid-pattern formation, and I establish a direct link to classical pattern-forming systems of the Turing type (Turing, 1952).

5.2. Model of grid-pattern formation

5.2.1. Model of neural activity

I consider a single cell that receives synaptic input from N spatially-tuned excitatory neurons (Figure 5.1A). Input spike trains $S_i^{\text{in}}(t) := \sum_k \delta(t - t_{i,k}^{\text{in}})$ for $i = 1, 2, \ldots, N$ are generated by independent inhomogeneous Poisson processes with instantaneous rates $r_i^{\text{in}}(t)$ where $\delta(t)$ is the Dirac delta function, and $t_{i,k}^{\text{in}}$ is the timing of the k^{th} input spike at synapse i. Similarly, the output spike train $S^{\text{out}}(t) := \sum_k \delta(t - t_k^{\text{out}})$ is generated by an inhomogeneous Poisson process with instantaneous rate $r^{\text{out}}(t)$ where t_k^{out} denotes the timing of the k^{th} output spike.

I assume that inputs are integrated linearly at the output, and that the output neuron is equipped with an intrinsic spike-rate adaptation mechanism, that is,

$$r^{\text{out}}(t) := r_0 + \int_0^\infty \mathrm{d}\tau \; K(\tau) \; \sum_{i=1}^{N} w_i \, S_i^{\text{in}}(t - \tau) \tag{5.1}$$

where r_0 is a baseline rate, w_i is the synaptic weight of input neuron i, and the function K is a temporal filter modeling the spike-rate adaptation dynamics. Note that the instantaneous output rate r^{out} depends only on the temporal history of the input spikes and that there is no reset mechanism after the emission of an output spike.

The adaptation kernel K is the sum of two exponential functions:

$$K(t) := \begin{cases} \dfrac{1}{\tau_{\text{S}}} \exp\left(-\dfrac{t}{\tau_{\text{S}}}\right) - \dfrac{\mu}{\tau_{\text{L}}} \exp\left(-\dfrac{t}{\tau_{\text{L}}}\right) & \text{for} \quad t \geq 0 \\ 0 & \text{for} \quad t < 0 \end{cases} \tag{5.2}$$

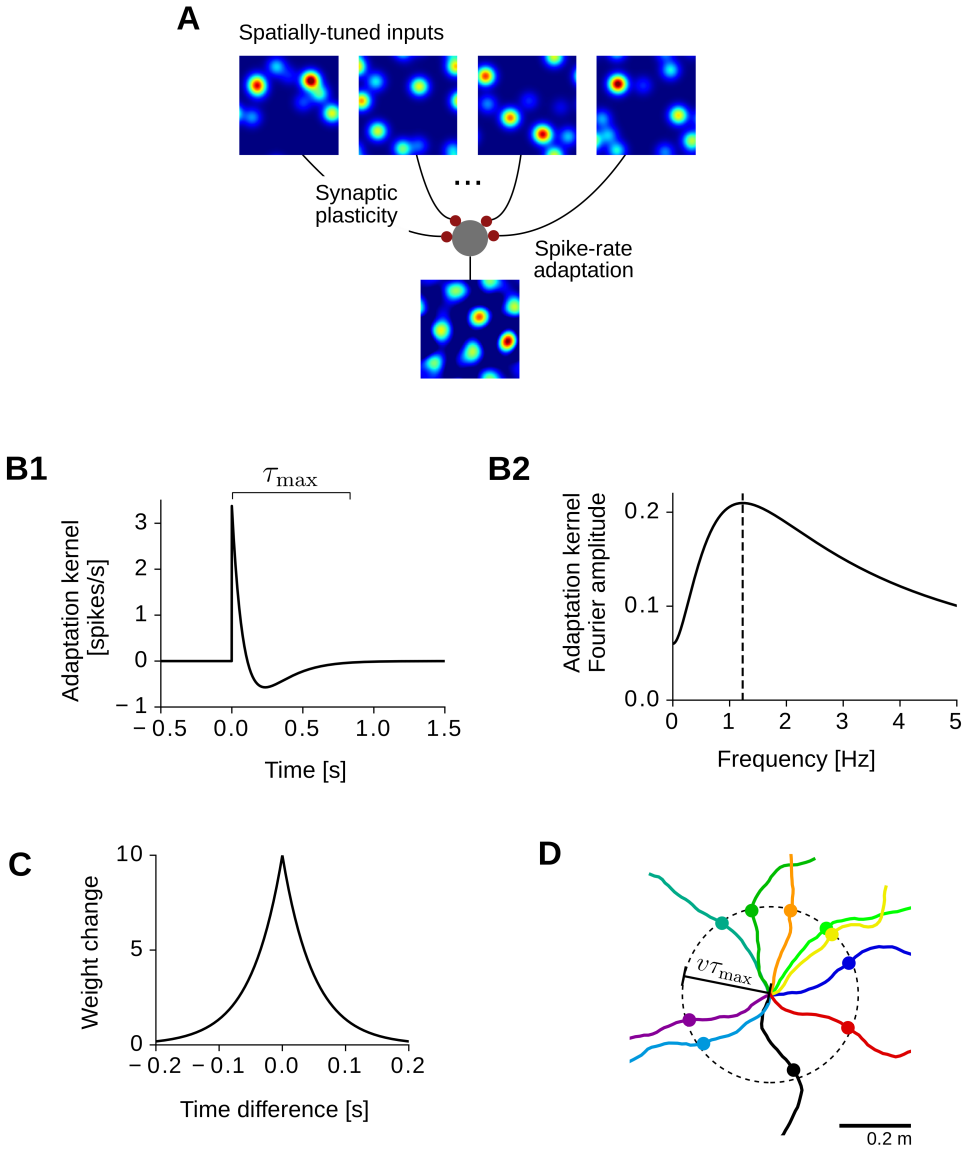

Figure 5.1. | **Model of grid-pattern formation**. **A)** Illustration of the model architecture. A set of spatially-tuned inputs (firing-rate maps at the top) project to a single cell (output neuron, gray disk). The input synapses are plastic and the output neuron shows spike-rate adaptation. After learning, the output cell develops a grid-like firing pattern (firing-rate map at the bottom). Warm colors in the firing-rate maps denote locations of high activity in the environment. **B)** Model of spike-rate adaptation. B1) The adaptation kernel K (Equation 5.2). A positive peak with amplitude $K(0) = 1/\tau_S - \mu/\tau_L \approx 3.4$ spikes/s is followed by a slow negative response. Note that the kernel is small for $t > \tau_{max}$, i.e., $|K(t)| < 0.01|K(0)|$ for $t > \tau_{max}$, with $\tau_{max} = 5\tau_L = 0.8$ s. B2) Frequency response of the adaptation kernel. The dashed vertical line indicates the kernel's resonance frequency $k_{res} = 1.23\,s^{-1}$. Parameter values: $\tau_S = 0.1$ s, $\tau_L = 0.16$ s, $\mu = 1.06$. The integral of the kernel is $1 - \mu = -0.06$. **C)** Learning window for spike-timing dependent plasticity (Equation 5.5). The horizontal axis denotes the time difference between input and output spikes. **D)** Model of spatial exploration. Colored lines denote example virtual-rat trajectories starting at the center of the black dashed circle (Equation 5.11). Filled dots indicate the position of the virtual rat at time $\tau_{max} = 5\tau_L = 0.8$ s. Movement trajectories are smooth within time stretches shorter than τ_{max}. Parameter values: $v = 0.25$ m/s, $\theta_\sigma = 0.7$. The circle radius is $v\tau_{max} = 20$ cm.

where τ_S and τ_L are the short and long kernel time constants ($0 < \tau_S < \tau_L$), and the parameter $\mu > 0$ sets the kernel integral $\int_0^\infty \mathrm{d}t\, K(t) = 1 - \mu$ (Figure 5.1B1). Intuitively, at the arrival of an input spike, the firing probability of the output neuron is first increased for a short time that is controlled by the time constant τ_S, and then decreased for a longer time that is controlled by the time constant τ_L. This second hyper-polarization dynamics effectively hinders the neuron to fire at high rates for long stretches of time, mimicking a spike-rate adaptation mechanism (Alonso and Klink, 1993; Van der Linden and da Silva, 1998; Yoshida et al., 2013). From a signal-processing perspective, the adaptation kernel K performs a temporal band-pass filtering of the input activity (Figure 5.1B2), and the two time constants τ_S and τ_L control the resonance frequency k_{res} at which the filter response is maximal. Note that in Section 5.A.5 I study a variant of the present model where neuronal adaptation is obtained through after-spike hyperpolarizing potentials associated to the output activity of the neuron.

5.2.2. Model of synaptic plasticity

I assume spike-timing dependent plasticity (STDP) at the input synapses (e.g., Gerstner et al., 1996; Markram et al., 1997; Kempter et al., 1999; Song et al., 2000; Zhou et al., 2005). Input and output spikes trigger weight changes Δw_i according to the following rule:

1. For each pair of a postsynaptic spike and a presynaptic spike at synapse i, I set

$$\Delta w_i = \eta W(\Delta t) \tag{5.3}$$

2. For each presynaptic spike at synapse i, I set

$$\Delta w_i = \eta(\beta - \alpha w_i) \tag{5.4}$$

where $\eta \ll 1$ is a small learning rate, and the STDP learning window $W(\Delta t)$ sets the weight change as a function of the time difference $\Delta t := t_{pre} - t_{post}$ between pre- and postsynaptic spikes. I consider a symmetric STDP learning window (Mishra et al., 2016, Figure 5.1C)

$$W(\Delta t) := \frac{W_{tot}}{2\tau_W} \exp(-|\Delta t|/\tau_W) \tag{5.5}$$

where the time constant $\tau_W > 0$ controls the maximal time lag at which plasticity occurs, and $W_{tot} = \int_{-\infty}^{\infty} \mathrm{d}t\, W(t)$ is the integral of the learning window. The first part of the learning rule (Equation 5.3) is the classical Hebbian term whereas the second part (Equation 5.4) is a local normalization term that stabilizes the average synaptic strength $w_{av} = N^{-1} \sum_{i=1}^{N} w_i$ and prevents the individual weights to grow

unbounded. This normalization term mimics local homeostatic processes observed experimentally (O'Brien et al., 1998; Davis and Goodman, 1998; Turrigiano et al., 1998); see also (Turrigiano and Nelson, 2004) for a review. The parameters $\alpha > 0$ and β set, respectively, the rate of weight decay and the target average weight w_{av} (Section 5.A.1). Importantly, the synaptic weights are constrained to non-negative values by imposing the hard bounds

$$w_i \geq 0 \ \forall i \,. \tag{5.6}$$

5.2.3. Model of input spatial tuning

I consider excitatory inputs with firing rates r_i^{in} that are tuned to the spatial position of a virtual rat exploring a square arena of side-length L, i.e.,

$$r_i^{\text{in}}(t) := \Psi_i^{\text{in}}(\mathbf{x}_t) \tag{5.7}$$

where $\mathbf{x}_t$ is the position of the virtual rat at time t, and Ψ_i^{in} is a spatial tuning curve. I characterize the spatial tuning curves Ψ_i^{in} in two alternative scenarios:

1. *spatially-regular* inputs, i.e., each input has a single spatial receptive field;
2. *spatially-irregular* inputs, i.e., each input has multiple spatial receptive fields at random locations (see examples in Figure 5.1A).

The first scenario, which is reminiscent of hippocampal place-cell activity (O'Keefe, 1976; Keefe and Burgess, 1996; Mizuseki et al., 2012), is easier to study analytically and cheaper to simulate numerically. The second scenario, which is reminiscent of parasubicular activity (Cacucci et al., 2004; Hargreaves et al., 2005, 2007; Boccara et al., 2010; Tang et al., 2016), is motivated by the anatomy of the entorhinal circuit (Section 2.1.1). In both cases, I consider circularly-symmetric receptive fields that cover the arena evenly. Indeed, place fields in open environments do not show systematic shape biases, and, in the absence of reward or goal locations, they are roughly homogeneously distributed (O'Keefe, 1976; Keefe and Burgess, 1996; Mizuseki et al., 2012; Cacucci et al., 2004; Hargreaves et al., 2005, 2007; Boccara et al., 2010; Tang et al., 2016). Note, however, that border-like inputs (Solstad et al., 2008; Lever et al., 2009)—which are not radially-symmetric—are present in the real system, but not explicitly modeled here. Finally, for simplicity, I assume periodic boundaries at the edges of the arena.

Spatially-regular inputs

In the case of *spatially-regular* inputs, I assume tuning curves of the form

$$\Psi_i^{\text{in}}(\mathbf{x}) := \mathcal{G}(|\mathbf{x} - \mathbf{r}_i|) \quad \text{with} \quad i = 1, 2, \ldots, N \tag{5.8}$$

where $\mathbf{r}_i$ is the receptive-field center of neuron i and $\mathcal{G}$ is a Gaussian function:

$$\mathcal{G}(r) := \frac{L^2 r_{\mathrm{av}}}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) . \tag{5.9}$$

The parameter $\sigma > 0$ sets the width of the receptive field, and r_{av} is the average firing rate in the environment. I assume that the input receptive-field centers $\mathbf{r}_i$ cover the entire arena evenly.

This input scenario is considered for the mathematical derivations in Section 5.3.2 and for the numerical simulations in Sections 5.4.1 and 5.4.2.

Spatially-irregular inputs

In the case of *spatially-irregular* inputs, each tuning curve Ψ_i^{in} is the sum of $M > 1$ Gaussian receptive fields with random amplitudes A_{ij} and random receptive-field centers $\mathbf{r}_{ij}$ with $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, M$, that is,

$$\Psi_i^{\mathrm{in}}(\mathbf{x}) := \frac{1}{\beta_i} \sum_{j=1}^{M} A_{ij}\, \mathcal{G}(|\mathbf{x} - \mathbf{r}_{ij}|) . \tag{5.10}$$

The scaling factors $\beta_i = \sum_{j=1}^{M} A_{ij}$ normalize the inputs Ψ_i^{in} to the same average rate r_{av}, and all the superimposed fields share the same field size σ (Equation 5.9). The field amplitudes A_{ij} are uniformly distributed in the range $(0, 1)$, and the receptive-field centers $\mathbf{r}_{ij}$ are uniformly distributed in the environment.

This input scenario is considered for the mathematical derivations in Section 5.A.4 and for the numerical simulations in Section 5.4.3.

5.2.4. Model of spatial exploration

The movement of the virtual rat follows a smooth random walk that satisfies the following three assumptions: (i) the movement speed v is constant in time; (ii) the random walk is isotropic and ergodic with respect to the auto-covariance; (iii) the virtual-rat trajectories are smooth within time stretches shorter than the time length $\tau_{\max} = 5\tau_{\mathrm{L}}$ of the adaptation kernel K (Figure 5.1D). Note that assumption (i) is obviously not valid in general. However, because synaptic plasticity acts on a time scale that is much slower than behavior, the relevant variable for pattern formation is the rat running speed averaged over long stretches of time (e.g. minutes), which can be considered approximately constant. I assume an average running speed of 25 cm/s, which is experimentally plausible (Reifenstein et al., 2014). Assumptions (ii) and (iii) hold by ignoring directional anisotropies deriving from the geometry of the environment, and by observing that experimental rat trajectories are approximately straight over short

running distances (e.g., over distances shorter than 25 cm; Reifenstein et al., 2014).

Mathematically, the two-dimensional virtual-rat trajectories $\mathbf{x}_t$ are sampled from the stochastic process

$$\frac{\mathrm{d}\mathbf{X}_t}{\mathrm{d}t} := v\,[\cos(\theta_t), \sin(\theta_t)] \quad \text{with} \quad \theta_t = \sigma_\theta \mathcal{W}_t\,, \tag{5.11}$$

where the angle θ_t sets the direction of motion and $\mathcal{W}_t$ is a standard Wiener process. The parameters v and σ_θ control the speed of motion and the tortuosity of the trajectory. Note that I also perform simulations with variable running speeds. In this case, the speed is sampled from an Ornstein-Uhlenbeck process with long-term mean $\bar{v} = v$.

5.3. Analytical results on grid-pattern formation

The grid-cell model presented in Section 5.2 is studied both analytically and numerically. In this section, I obtain an equation for the average dynamics of the synaptic weights, and I derive analytical requirements for spatial pattern formation. In Section 5.4, I demonstrate the emergence of grid-like activity by simulating both the detailed spiking model and the averaged system. The results presented here may be skipped by the less mathematically-inclined reader.

5.3.1. Average weight dynamics

I study structure formation in the activity of an output cell by averaging the weight dynamics resulting from the stochastic activation of input and output neurons (Section 5.2.1) and the STDP learning rule (Section 5.2.2), while a virtual rat explores a two-dimensional enclosure and the inputs are spatially tuned (Sections 5.2.3 and 5.2.4). I take both ensemble averages across spike-train realizations and temporal averages within a time window of length T. The averaging time length T separates the time scale of neural activation (of the order of the width τ_W of the learning window W) from the time scale τ_{str} of structure formation, i.e., $\tau_W \ll T \ll \tau_{str}$. Because τ_{str} is inversely proportional to the learning rate η (Equation 5.29), such averaging is always possible provided that the learning rate η is small enough. In other words, I assume that within a time T, the virtual rat has roughly explored the entire environment, but the synaptic weights did not change considerably. In this case, the dynamics of the synaptic weights w_i is approximated by a drift-diffusion process, where the deterministic drift term reads (Kempter et al., 1999)

$$\eta^{-1}\frac{\mathrm{d}\bar{w}_i}{\mathrm{d}t} = (\beta - \alpha\bar{w}_i)\overline{\langle S_i^{\mathrm{in}}(t)\rangle} + \int_{-\infty}^{\infty} \mathrm{d}s\, W(s)\overline{\langle S_i^{\mathrm{in}}(t+s)S^{\mathrm{out}}(t)\rangle} \tag{5.12}$$

with $\bar{w}_i \geq 0$. The functions S_i^{in} and S^{out} denote input and output spike trains (Section 5.2.1), the angular brackets denote ensemble averages over input and output spike trains, and the overbars denote temporal averages, i.e., $\bar{f}(t) := T^{-1} \int_{t-T}^{t} \mathrm{d}s f(s)$. Following Kempter et al. (1999), I derive

$$\overline{\langle S_i^{\text{in}}(t+s) S^{\text{out}}(t) \rangle} = \overline{\langle S_i^{\text{in}}(t+s) \rangle \langle S^{\text{out}}(t) \rangle} + \bar{w}_i \, \overline{\langle S_i^{\text{in}}(t) \rangle} K(-s) \,, \tag{5.13}$$

where the ensemble averages read

$$\langle S_i^{\text{in}}(t) \rangle = r_i^{\text{in}}(t) \tag{5.14}$$

$$\langle S^{\text{out}}(t) \rangle = \langle r^{\text{out}}(t) \rangle \overset{(5.1)}{=} r_0 + \int_0^\infty \mathrm{d}\tau \; K(\tau) \; \sum_{j=1}^{N} w_j r_j^{\text{in}}(t-\tau) \,. \tag{5.15}$$

Finally, from Equations 5.12-5.15, I obtain

$$\eta^{-1} \frac{\mathrm{d}}{\mathrm{d}t} \bar{w}_i = \sum_{j=1}^{N} C_{ij} \bar{w}_j - a \bar{w}_i + b \quad \text{with} \quad \bar{w}_i \geq 0 \tag{5.16}$$

where I defined

$$C_{ij} := \int_0^\infty \mathrm{d}\tau \, K(\tau) \int_{-\infty}^{\infty} \mathrm{d}s \, W(s) \, \overline{r_i^{\text{in}}(t+s) r_j^{\text{in}}(t-\tau)} \tag{5.17}$$

$$a := r_{\text{av}} \left[\alpha - \int_{-\infty}^{\infty} \mathrm{d}s \, W(s) K(-s) \right] \tag{5.18}$$

$$b := r_{\text{av}} \left(W_{\text{tot}} r_0 + \beta \right) \,. \tag{5.19}$$

Note that in deriving Equation 5.16 I approximated the temporal average of the input rates $\overline{r_i^{\text{in}}}$ with the spatial average r_{av} of the input tuning curves Ψ_i^{in}. This approximation holds with the assumption that in a time T the virtual rat roughly covers the entire space evenly.

By ignoring the non-linear weight constraints $\bar{w}_i \geq 0$, the average weight dynamics is described by a linear system with coupling terms C_{ij} (Equation 5.16). The coefficients C_{ij} are given by the temporal correlations of the input rates r_i^{in} and r_j^{in}, filtered by the adaptation kernel K and the STDP learning window W (Equation 5.17).

To further simplify the calculations, I assume that the low-pass filtering introduced by the STDP learning window can be neglected for the purpose of studying pattern formation. In particular, I assume that the learning window W decays much faster than the changes in the input correlations $\overline{r_i^{\text{in}}(t+s) r_j^{\text{in}}(t-\tau)}$ (Equation 5.17), which holds for $\tau_{\text{W}} \ll \sigma / v$. In this case, I obtain

$$C_{ij} \approx W_{\text{tot}} \int_0^\infty \mathrm{d}\tau \; K(\tau) \; \overline{r_i^{\text{in}}(t) r_j^{\text{in}}(t-\tau)} \tag{5.20}$$

where W_{tot} is the integral of the learning window (Equation 5.5).

Finally, by assuming smooth virtual-rat trajectories at constant speed v, the correlation matrix C_{ij} can be estimated solely from the input tuning curves Ψ_i^{in} and the adaptation kernel K (Section 5.A.2, Equation 5.47):

$$C_{ij} \approx \frac{W_{\text{tot}}}{L^2} \int_0^\infty \mathrm{d}\tau\, K(\tau) \oint_{|\mathbf{z}|=\tau v} \mathrm{d}\mathbf{z}\; \Psi_i^{\text{in}} \star \Psi_j^{\text{in}} \Big|_{\mathbf{z}} \tag{5.21}$$

where L^2 is the area explored by the virtual rat. In Equation 5.21, the matrix element C_{ij} is obtained by integrating the spatial cross-correlation of the input tuning curves $\Psi_i^{\text{in}} \star \Psi_j^{\text{in}}$ over circles of radius τv, and by weighting each integral with the amplitude of the adaptation kernel K at time τ. Note that Equation 5.21 holds for generic spatial tuning curves Ψ_i^{in}.

5.3.2. Weight dynamics for spatially-regular inputs

To study the emergence of spatial patterns, I now consider the simplified scenario of spatially-regular inputs (Section 5.2.3). That is, the input tuning curves Ψ_i^{in} are circularly-symmetric Gaussian functions that cover the entire space evenly (Equations 5.8 and 5.9). This input representation is particularly useful because it establishes a direct mapping between the neuron identity (the index i) and a position in physical space (the receptive-field center $\mathbf{r}_i$). Therefore, studying pattern formation in the activity of the output neuron is reduced to studying pattern formation in the space of the synaptic weights. Note, however, that such a simple input scenario is not necessary for the emergence of grid patterns in general, as shown in Section 5.4.3.

With spatially-regular inputs, the average weight dynamics in Equation 5.16 can be rewritten by labeling the synaptic weights according to the corresponding receptive-field centers $\mathbf{r}_i$:

$$\eta^{-1} \frac{\mathrm{d}}{\mathrm{d}t} \bar{w}(\mathbf{r}_i) = \sum_{i=1}^{N} C(\mathbf{r}_i, \mathbf{r}_j) \bar{w}(\mathbf{r}_j) - a\bar{w}(\mathbf{r}_i) + b \tag{5.22}$$

where $\bar{w}(\mathbf{r}_i) = \bar{w}_i$ and $C(\mathbf{r}_i, \mathbf{r}_j) = C_{ij}$. Additionally, in the limit of a large number $N \gg 1$ of input neurons and receptive fields that cover the environment with constant density $\rho = N/L^2$, the sum in Equation 5.22 can be replaced by an integral over all the receptive-field centers $\mathbf{r}'$:

$$\eta^{-1} \frac{\mathrm{d}}{\mathrm{d}t} \bar{w}(\mathbf{r}) = \rho \int \mathrm{d}\mathbf{r}'\, C(\mathbf{r}, \mathbf{r}') \bar{w}(\mathbf{r}') - a\bar{w}(\mathbf{r}) + b \tag{5.23}$$

where the correlation function $C(\mathbf{r}, \mathbf{r}')$ is the continuous extension of the correlation matrix $C_{ij} = C(\mathbf{r}_i, \mathbf{r}_j)$. Because the inputs are translation invariant (Equation 5.8), the correlation function C is also translation invariant, i.e., $C(\mathbf{r}, \mathbf{r}') = C(\mathbf{r} - \mathbf{r}', \mathbf{0}) = C(\mathbf{r} - \mathbf{r}')$,

where I omit the second argument $\mathbf{0} := (0,0)$ for readability. In this case, the integral in Equation 5.23 can be expressed as a two-dimensional convolution in space:

$$\eta^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\bar{w}(\mathbf{r}) = \rho\int \mathrm{d}\mathbf{r}'\, C(\mathbf{r}-\mathbf{r}')\bar{w}(\mathbf{r}') - a\bar{w}(\mathbf{r}) + b\,. \tag{5.24}$$

Figure 5.2 shows the correlation C as a function of the input receptive-field distance $|\mathbf{r}-\mathbf{r}'|$ for the adaptation kernel K in Figure 5.1B1 and Gaussian input fields with size $\sigma = 6.25$ cm (Equation 5.9). The function C has the shape of a typical Mexican-hat kernel, i.e., it is positive for short receptive-field distances (attraction domain), negative for intermediate distances (repulsion domain), and zero otherwise. In this case, the synaptic weights of close-by input fields grow together whereas the synaptic weights of input fields that are further apart are repelled from each other (Equation 5.24). Such a competitive Mexican-hat interaction is at the basis of many pattern-forming systems found in nature, and it is directly related to diffusion-driven instabilities of the Turing type (see, e.g., Murray, 2002).

5.3.3. Eigenvalue spectrum for spatially-regular inputs

To study spatially-periodic solutions, I take the two-dimensional Fourier transform with respect to $\mathbf{r}$ at both sides of Equation 5.24:

$$\eta^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\widehat{w}(\mathbf{k}) = (\rho\widehat{C}(\mathbf{k}) - a)\,\widehat{w}(\mathbf{k}) + \delta(\mathbf{k})b \tag{5.25}$$

where I defined the Fourier transform pair

$$\widehat{w}(\mathbf{k}) := \int \mathrm{d}\mathbf{r}\,\bar{w}(\mathbf{r})\exp(-2\pi j\,\mathbf{k}\cdot\mathbf{r})\ ,\ \bar{w}(\mathbf{r}) = \frac{1}{(2\pi)^2}\int \mathrm{d}\mathbf{k}\,\widehat{w}(\mathbf{k})\exp(2\pi j\,\mathbf{k}\cdot\mathbf{r})\,, \tag{5.26}$$

$\mathbf{k}$ is a two-dimensional wave vector, and $j = \sqrt{-1}$ is the imaginary unit. Solving Equation 5.25 for $\mathbf{k} \neq (0,0)$, I obtain

$$\widehat{w}(\mathbf{k}) = \widehat{w}_0(\mathbf{k})\exp(\eta\lambda(\mathbf{k})t) \tag{5.27}$$

where $\widehat{w}_0(\mathbf{k})$ denotes the weight spectrum at time $t = 0$, and I defined

$$\lambda(\mathbf{k}) := \rho\widehat{C}(\mathbf{k}) - a \quad \text{for} \quad \mathbf{k} \neq (0,0)\,. \tag{5.28}$$

The function $\lambda(\mathbf{k})$ defines the eigenvalue spectrum of the dynamical system in Equation 5.24, and the corresponding eigenfunctions are the elements of the Fourier basis $\exp(2\pi j\,\mathbf{k}\cdot\mathbf{r})$. Equation 5.28 is also called the *dispersion relation* of the system. Note that solving Equation 5.25 for $\mathbf{k} = (0,0)$ one obtains the dynamics of the total synaptic

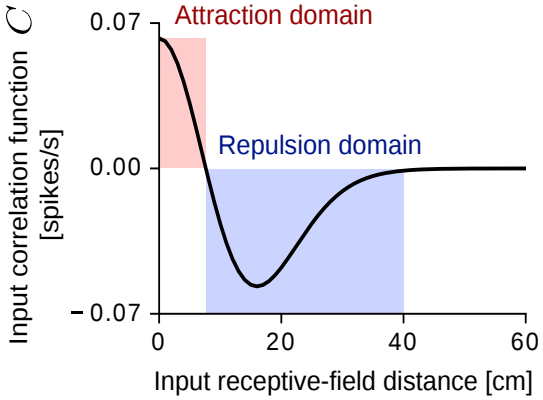

Figure 5.2. | **Input correlation function C for spatially-regular inputs.** The function is circularly symmetric, i.e., it depends only on the distance $|\mathbf{r} - \mathbf{r}'|$ between the receptive-field centers $\mathbf{r}$ and $\mathbf{r}'$ (Equation 5.54). In the attraction domain (red shaded area), the correlation is positive and the synaptic weights grow in the same direction. In the repulsion domain (blue shaded area), the correlation is negative and the synaptic weights grow in opposite directions. Parameter values: $\sigma = 6.25$ cm, $r_{\text{av}} = 0.4\ \text{s}^{-1}$, $\tau_{\text{S}} = 0.1$ s, $\tau_{\text{L}} = 0.16$ s, $\mu = 1.06$, $W_{\text{tot}} = 1$ s, $L = 1$ m, $v = 0.25$ m/s

weight, which is kept normalized by the learning rule (Section 5.A.1).

From Equation 5.27, the Fourier modes of the synaptic weights $\widehat{w}(\mathbf{k})$ grow or decay exponentially with rates proportional to the eigenvalues $\lambda(\mathbf{k})$. Therefore, a structure in the synaptic weights emerges on a time scale

$$\tau_{\text{str}} := \frac{1}{\eta \lambda_{\max}} \tag{5.29}$$

where $\lambda_{\max} := \max_{\mathbf{k}}[\lambda(\mathbf{k})]$ is the largest eigenvalue in the system.

Importantly, the eigenvalues $\lambda(\mathbf{k})$ are linearly related to the Fourier transform of the input-correlation function $\widehat{C}(\mathbf{k})$ (Equation 5.28), which is circularly-symmetric for circularly-symmetric inputs. In this case, in Section 5.A.3 (Equation 5.62) I derive

$$\widehat{C}(\mathbf{k}) \approx \frac{W_{\text{tot}}}{L^2} 4\pi^2 \tilde{\mathcal{G}}^2(k)\, \tilde{K}_{\text{sp}}(k) \quad \text{with} \quad k := |\mathbf{k}| \tag{5.30}$$

where $\tilde{\mathcal{G}}$ and $\tilde{K}_{\text{sp}}$ (Equations 5.63 and 5.64) are the zeroth-order Hankel transforms (Equation 5.59) of the input tuning curve $\mathcal{G}$ (Equation 5.8) and of the equivalent adaptation kernel in space

$$K_{\text{sp}}(r) \overset{(5.61)}{:=} \frac{1}{rv} K\left(\frac{r}{v}\right). \tag{5.31}$$

Finally, by plugging Equation 5.30 into Equation 5.28, I obtain

$$\lambda(\mathbf{k}) \approx \rho \frac{W_{\text{tot}}}{L^2} 4\pi^2 \tilde{\mathcal{G}}^2(k)\, \tilde{K}_{\text{sp}}(k) - a \quad \text{with} \quad k \neq 0\,. \tag{5.32}$$

From Equations 5.27 and 5.32, I recognize a necessary condition for spatial patterns to emerge: the eigenvalue spectrum $\lambda(\mathbf{k}) = \lambda(k)$ shall have a global maximum at a frequency $k_{\max} > 0$ with $\lambda_{\max} = \lambda(k_{\max}) > 0$. In this case, all the Fourier modes $\mathbf{k}$ at the critical frequency $|\mathbf{k}| = k_{\max}$ are unstable (Equation 5.25), and spatially-periodic patterns could emerge.

Figure 5.3 shows the critical frequency $k_{\max}$ (panels A1 and B1) and the largest eigenvalue $\lambda_{\max}$ (panels A2 and B2) as a function of the parameters of the adaptation kernel K, i.e., the short time constant τ_S, the long time constant τ_L, and the kernel integral $1 - \mu$ (Equation 5.2). The input receptive-field width σ is kept constant. In panels A1 and B1, the green-shaded regions correspond to parameter values where periodic grid-like patterns could emerge ($k_{\max} > 0$). Conversely, the white regions denote parameter values where place-cell-like receptive fields could emerge ($k_{\max} = 0$) (D'Albis et al., 2015). Note that the spatial scale of the periodic patterns depends on the long adaptation time constant τ_L (panel A1), but is largely unaffected by the short time constant τ_S (panel B1). Additionally, the largest spatial frequencies are obtained for small values of τ_L and negative kernel integrals (panel A1). This leads me to the following predictions: the grid scale shall depend on the long temporal dynamics of the adaptation kernel, and the smallest grid scales require adaptation kernels with an overall inhibitory effect on the activity of the output neuron. Also note that larger values of τ_L correspond to larger values of $\lambda_{\max}$ (panel A2). Thus, I predict that grids at larger scales shall develop faster than grids at smaller scales (Equation 5.29).

Importantly, the formation of grid-like patterns also requires a nonlinearity in the system. Indeed, for triangular lattices to emerge, only three wave vectors $\mathbf{k}$ of the same length $|\mathbf{k}|$ shall survive. But this cannot be achieved in a linear system where all Fourier modes develop independently from each other (Equation 5.27). Yet the non-linear weight constraints imposed in the model (Equation 5.6) are sufficient to generate triangular patterns (Section 5.4.1).

In summary, the theory presented here gives necessary conditions for spatial pattern formation, and it predicts how the shape of the adaptation kernel influences the scale of the grids and the relative time required for their formation. The theory remains however agnostic about the specific two-dimensional periodicity of the resulting patterns, i.e., it cannot predict whether the final solutions are, e.g., planar waves, square, rhomboidal, or triangular lattices. Further mathematical insights on this topic could be obtained by using perturbation methods (see e.g. Ermentrout and Cowan, 1979), but this is beyond the scope of the present study.

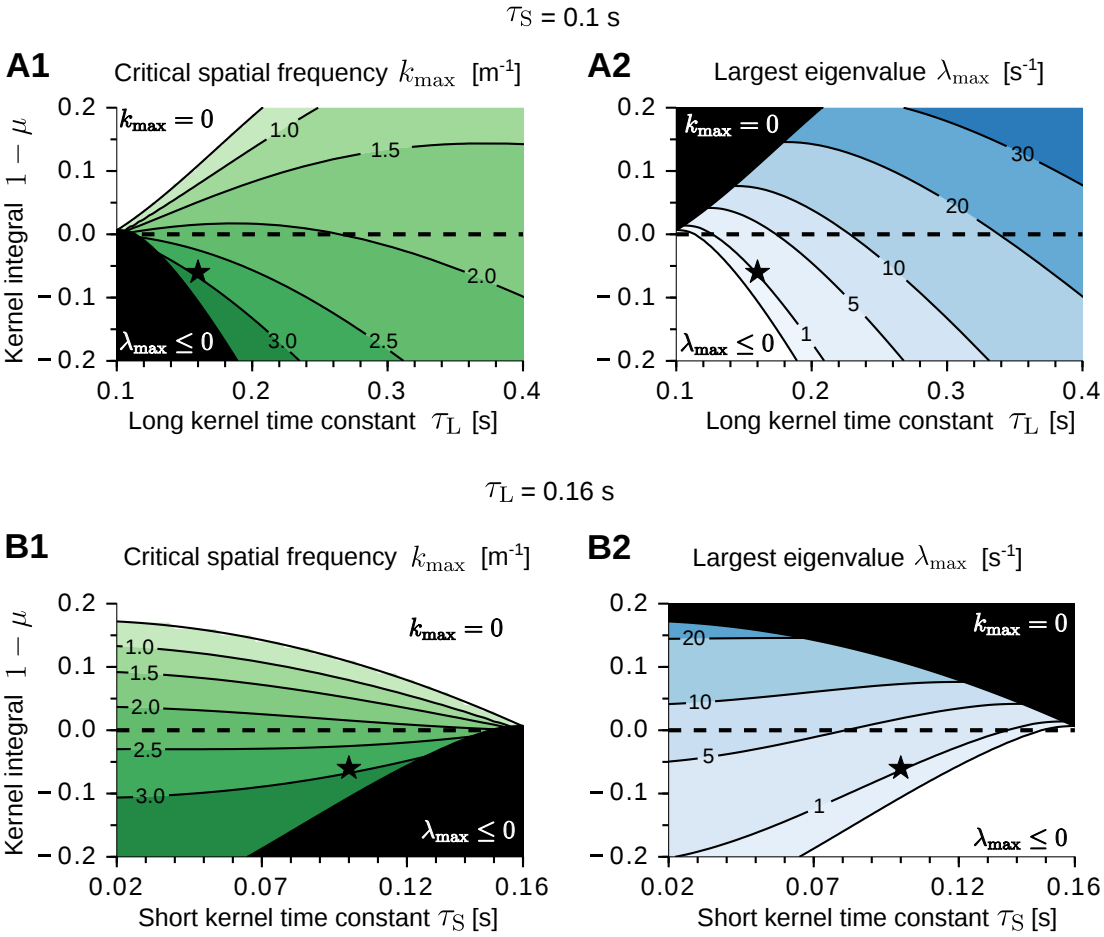

Figure 5.3. | Impact of the adaptation kernel on grid-pattern formation. A) Critical spatial frequency $k_{\max}$ (A1) and largest eigenvalue $\lambda_{\max}$ (A2) as a function of the kernel integral $1-\mu$ and the long kernel time constant τ_{L}. The short time constant is $\tau_{\mathrm{S}} = 0.1$ s. The black lines are iso-levels (see annotated values). Regions enclosed by two adjacent iso-lines are colored uniformly (darker colors denote larger values). Within the black region in A1, I obtain $\lambda_{\max} \leq 0\,\mathrm{s}^{-1}$ (see white region in A2). Within the black region in A2, I obtain $k_{\max} = 0\,\mathrm{m}^{-1}$ (see white region in A1). The dashed horizontal line indicates zero-integral kernels. The star denotes the parameter values $\tau_{\mathrm{S}} = 0.1$ s, $\tau_{\mathrm{L}} = 0.16$ s, $\mu = 1.06$ of the kernel in Figure 5.1B1. **B)** Same as in A, but varying the short kernel time constant τ_{S}. The long time constant is $\tau_{\mathrm{L}} = 0.16$ s. The eigenvalue spectrum is estimated from Equation 5.32. Further parameter values: $\sigma = 6.25$ cm, $r_{\mathrm{av}} = 0.4\,\mathrm{s}^{-1}$, $W_{\mathrm{tot}} = 1$ s, $\rho = 900\,\mathrm{m}^{-2}$, $L = 1$ m, $v = 0.25$ m/s, $a = 1.1\,\mathrm{s}^{-1}$.

5.4. Numerical results on grid-pattern formation

In Section 5.3, I derived an equation for the average dynamics of the synaptic weights w_i, under the STDP learning rule and the stochastic activation of input and output neurons (Equation 5.16). In the case of spatially-regular inputs, I then computed the systems eigenvalue spectrum $\lambda(k)$ in terms of the Gaussian input tuning curve $\mathcal{G}$ and the temporal adaptation kernel K (Equation 5.32). I showed that periodic spatial patterns could emerge if the eigenvalue spectrum $\lambda(k)$ had a global maximum at a frequency $k_{\max} > 0$ with $\lambda(k_{\max}) > 0$ (Figure 5.3).

Figure 5.4A shows the eigenvalue spectrum $\lambda(k)$ for a choice of the parameter values such that this condition is satisfied. With adaptation time constants $\tau_S = 0.1$ s and $\tau_L = 0.16$ s (Equation 5.2, Figure 5.1B1, star in Figure 5.3), and Gaussian input receptive fields of size $\sigma = 6.25$ cm (Equation 5.9), the eigenvalue spectrum peaks at the critical frequency $k_{\max} = 3\ \mathrm{m}^{-1}$. The evolution of the synaptic weights is simulated in this scenario.

5.4.1. Emergence of grid spatial patterns

First, I simulate the detailed spiking model with spatially-regular inputs (Section 5.2.3). The results are shown in Figure 5.4B-E. In line with the theory, a structure emerges in the synaptic weights (Figure 5.4B and C) on a time scale of $\tau_{\mathrm{str}} = 1/(\eta\lambda_{\max}) \approx 5 \cdot 10^4$ s (Equation 5.29) where $\eta = 2 \cdot 10^{-5}$ is the learning rate and $\lambda_{\max} \approx 1\ \mathrm{s}^{-1}$ is the largest eigenvalue in the system. Additionally, the weight spectrum is quickly dominated by the critical frequency $k_{\max} = 3\ \mathrm{m}^{-1}$ (Figure 5.4C, bottom row) at which the eigenvalue spectrum has a global maximum (Figure 5.4A).

Importantly, the synaptic weights also develop a periodic triangular symmetry, which is reminiscent of grid-cell patterns. Such triangular symmetry emerges after a substantial fraction of weights has hit the low saturation bound (Equation 5.6, Figure 5.4B, inset). Periodic pattern formation is indeed a strictly non-linear phenomenon, and excluding the spike generation process, weight saturation is the only non-linearity present in the system. In the linear regime, all Fourier modes $\mathbf{k}$ with frequency $|\mathbf{k}| = k_{\max}$ exponentially grow with equal rate $\eta\lambda(k_{\max})$ and *independently* from each other (Equation 5.25). In this case, the random weight pattern at time $t = 0$ s is amplified at the frequency $k_{\max}$, but no periodic structure emerges. In the non-linear regime, instead, the exponentially growing modes are mutually coupled, and a spontaneous symmetry breaking occurs: only three Fourier modes with wave vectors that are 60 degrees apart survive in my simulations (see Figure 5.4C and D).

In the example of Figure 5.4, a triangular symmetry starts to emerge after $2 \cdot 10^5$ s, i.e, about 50 hours of exploration of the virtual rat. Yet the time scale of learning depends

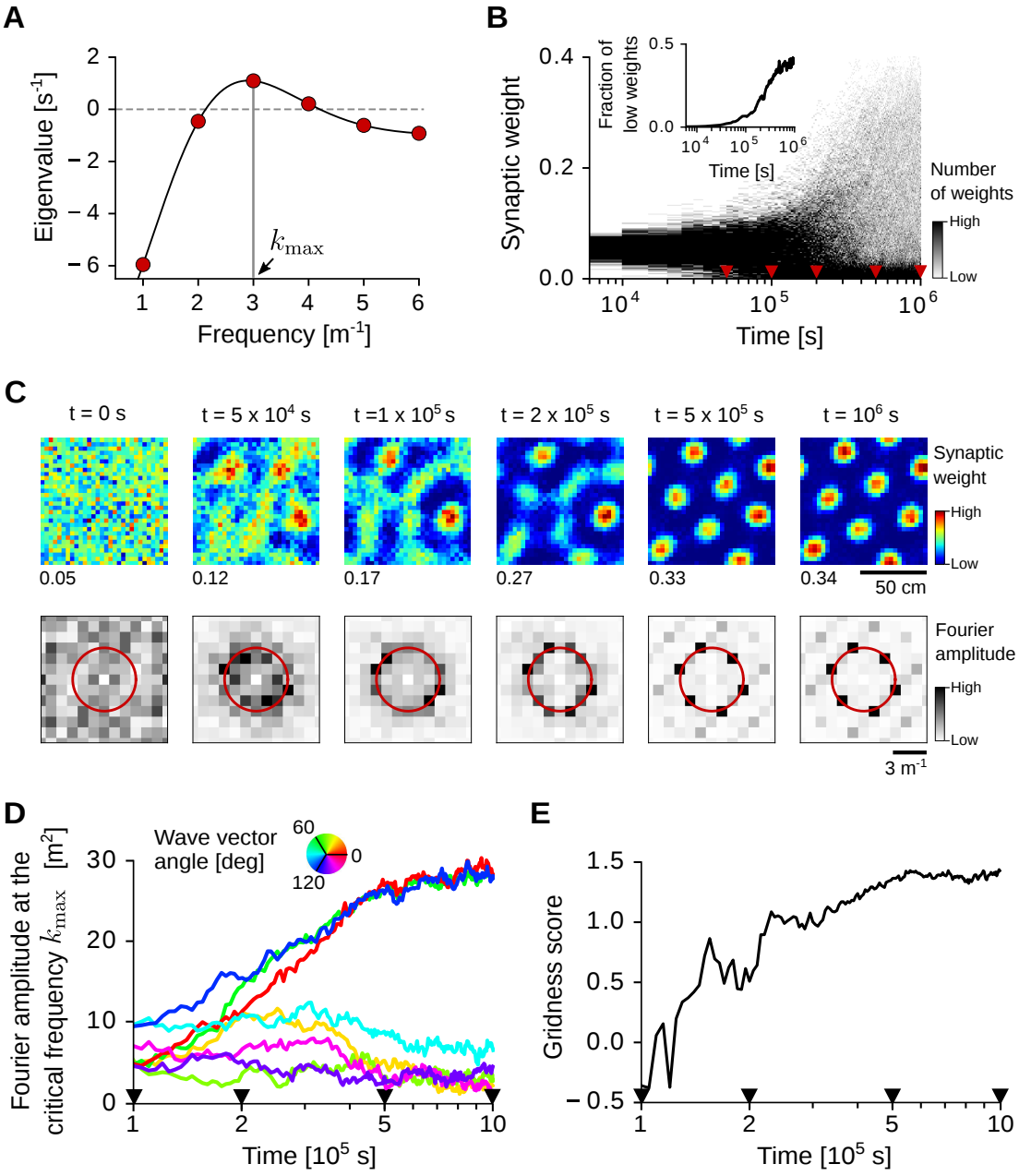

Figure 5.4. | Grid-pattern formation with spatially-regular inputs. **A)** Eigenvalue spectrum $\lambda(k)$ of the averaged weight dynamics (Equation 5.32). The black solid line shows the continuous spectrum in the limit of infinite-size environments; the red dots show the discrete eigenvalues for a square arena of side length $L = 1$ m with periodic boundaries. The horizontal dashed line separates positive and negative eigenvalues. The vertical gray line indicates the critical spatial frequency $k_{\max} = 3$ m^{-1}. The eigenvalue at frequency $k = 0$ is not shown. Parameter values: $\tau_S = 0.1$ s, $\tau_L = 0.16$ s, $\sigma = 6.25$ cm. **B)** Time-resolved distribution of $N = 900$ synaptic weights updated according to the STDP rule in Equations 5.3-5.6. Red triangles indicate the time points shown in C. Inset: fraction of weights close to the lower saturation bound ($w_i < 5 \cdot 10^{-3}$). **C)** Top row: evolution of the synaptic weights over time. Weights are sorted according to the two-dimensional position of the corresponding input receptive-field centers. Note that each panel has a different color scale (maximum weight at the bottom-left corner, see B for distributions). Bottom row: Fourier amplitude of the synaptic weights at the top row. The red circle indicates the frequency $k_{\max} = 3$ m^{-1} of the largest eigenvalue (see panel A). **D)** Time evolution of weights' Fourier amplitudes $|\widehat{w}(\mathbf{k})|$ for wave vectors $\mathbf{k}$ at the critical frequency $|\mathbf{k}| = k_{\max}$. Wave vector angles (color coded) are relative to the largest mode at the end of the simulation ($t = 10^6$ s). The black triangles indicate time points in C. **E)** Gridness score of the weight pattern over time. The gridness score quantifies the degree of triangular periodicity. See Section 5.A.6 for further details and parameter values.

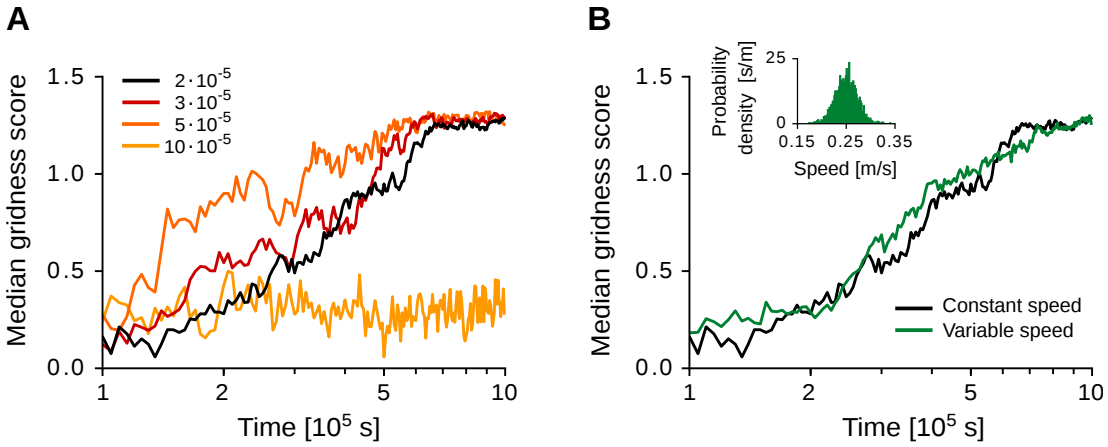

Figure 5.5. | **Time scales of learning**. (A) Median gridness scores of the input synaptic weights for 40 random weight initializations and different learning-rate values, i.e., $\eta = (2, 3, 5, 10) \cdot 10^{-5}$. The weight development is simulated with the detailed spiking model with spatially-regular inputs and constant virtual-rat speed (see also Figure 5.4). (B) Median gridness scores of the input synaptic weights simulated with constant (black line) and variable (green line) virtual-rat speeds for 40 random weight initializations. Variable running speeds are obtained by sampling from an Ornstein-Uhlenbeck process with long-term mean $\bar{v} = 0.25$ m/s, volatility $\sigma_v = 0.1$ m·s$^{-1.5}$ and mean-reversion speed $\theta_v = 10$ s^{-1}. The inset shows the distribution of running speeds (mean: 0.25 m/s std: 0.02 m/s). Note that the long-term mean $\bar{v}$ of the process equals the speed v in constant-speed simulations. See Section 5.A.6 for further details and additional parameter values.

on the learning rate η and on the largest eigenvalue $\lambda_{\max}$ (Equation 5.29), which are under-constrained by experimental data (see Equation 5.32 for the dependence of $\lambda_{\max}$ on other model parameters). From a theoretical standpoint, the speed of learning is limited by the noise in the system, which is due to the virtual-rat random walk and the stochastic spiking of the neurons. To theoretically explore this limit and test the robustness of the model against noisy initial conditions, I simulated the development of the synaptic weights for different values of the learning rate and multiple random initializations of the synaptic weights. The results are reported in Figure 5.5A. With larger learning rates, grid-like patterns emerge faster. However, if the learning rate is too large, e.g., $\eta = 10 \cdot 10^{-5}$ in my simulations, the gridness score fluctuates at low levels and no stable grid pattern emerges (yellow line in Figure 5.5A). Therefore, my results suggest that tens of hours of spatial exploration are required for stable grid patterns to emerge. Finally, the model is robust to random initializations of the synaptic weights, and to variations of the running speed of the virtual rat (Figure 5.5B).

5.4.2. Geometrical properties of the grid patterns

I now discuss the geometrical properties of the simulated grid patterns. A periodic triangular grid is characterized by three fundamental properties: i) the grid scale, i.e., the distance between two neighboring peaks; ii) the grid spatial phase, i.e., the spatial offset of the grid peaks with respect to a reference point; and iii) the grid orientation, i.e, the angle between one of the three grid axes and a reference direction (Section 3.1.1, Figure 3.1B).

Grid scale

In my model, the grid scale is set by the critical frequency $k_{\max}$ at which the eigenvalue spectrum has a global maximum (Equation 5.32 and Figure 5.4). This critical frequency depends only on the movement speed v of the virtual rat, the width σ of the input tuning curve $\mathcal{G}$, and the temporal dynamics of the adaptation kernel K (Figure 5.3). Therefore, grid patterns at different scales are obtained, for example, by varying the width σ of the input receptive fields or the long time scale τ_{L} of the adaptation kernel (Figure 5.6, see also Figure 5.3). This theoretical result is consistent with the facts that spatial tuning in the hippocampal formation is typically broader ventrally than dorsally (Jung et al., 1994; Fyhn et al., 2004; Kjelstrup et al., 2008), and that grid scales vary in the same direction (Hafting et al., 2005; Stensola et al., 2012). Additionally, I predict that the adaptation time scale may also have a dorso-ventral gradient, similarly to other intrinsic cellular properties in the MEC (e.g., Giocomo et al., 2007; Giocomo and Hasselmo, 2008b; Garden et al., 2008; Pastoll et al., 2012; Yoshida et al., 2013, Section 2.2.2).

Grid spatial phase

With evenly-distributed input fields and periodic boundaries, the spatial phases of the grid patterns depend only on the initial condition of the synaptic weights, i.e., random weight initializations result in uniformly-distributed grid phases (Figure 5.7A1 and B1). This result is in line with the phases of nearby grid cells being roughly evenly distributed in experimental data (Hafting et al., 2005), but see also (Heys et al., 2014). Yet it remains unclear whether the same results would be obtained in the case of non-periodic boundaries.

Grid orientation

With periodic boundary conditions, the model produces grid orientations that are distributed non-uniformly. Precisely, the distribution of grid orientations depends on the scale of the pattern relative to the size of the environment, e.g., in the same

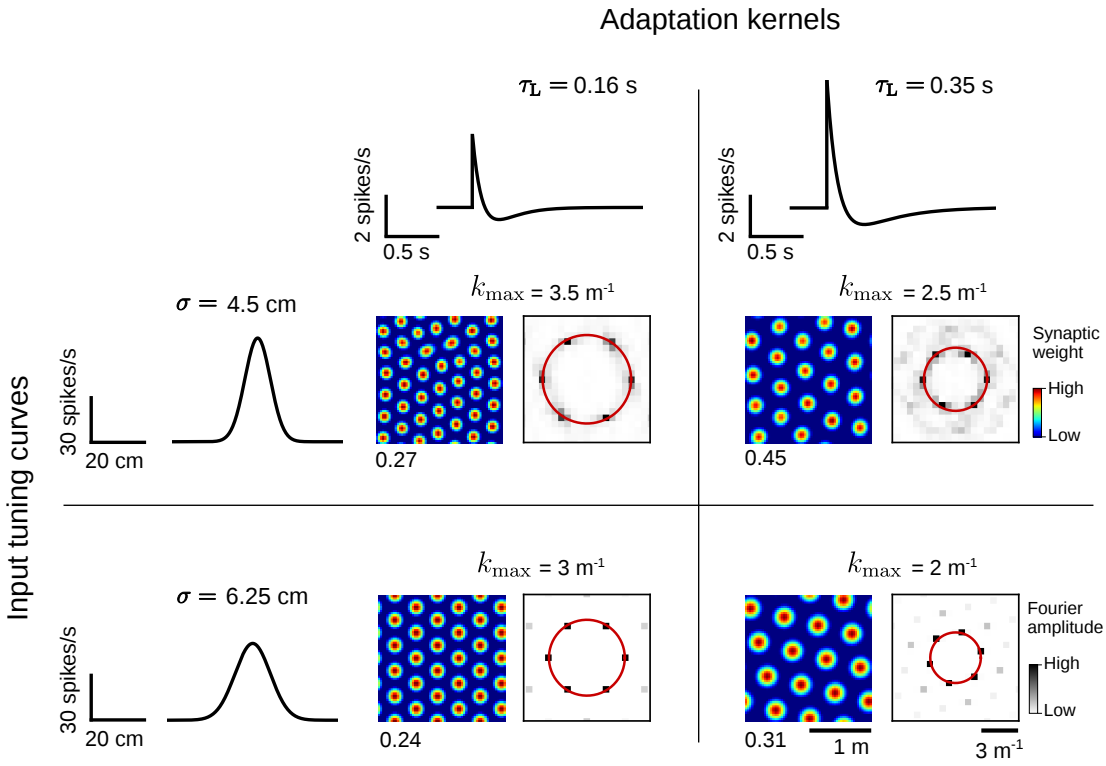

Figure 5.6. | Spatial scale of the grid patterns. Example grid patterns obtained with different adaptation kernels K (Equation 5.2, top row) and different input tuning curves $\mathcal{G}$ (Equation 5.9, left-most column). For each choice of the functions K and $\mathcal{G}$, the synaptic weights (left) and their corresponding Fourier spectra (right) at the end of the simulation are shown ($t = 10^6$ s). The synaptic-weight maps have different color scales (maximal values at the bottom-left corner). The red circles indicate the spatial frequency k_{max} of the weight patterns. Synaptic weights were obtained by simulating the average weight dynamics in Equation 5.16. Note that I used a larger enclosure ($L = 2$ m) as compared to the one in Figure 5.4 ($L = 1$ m). See Section 5.A.6 for further details and parameter values.

environment patterns at different scales tend to align differently (compare panels A2 and B2 in Figure 5.7). In the examples of Figure 5.7A3, one of the grid axes tends to align to a border of the arena whereas in the examples of Figure 5.7B3 one of the grid axes tends to align to a diagonal of the arena. Similar results are obtained by keeping the grid scale fixed and varying the size of the environment, e.g., compare Figure 5.7A2 ($k_{max} = 3$ m^{-1} and $L = 2$ m) and Figure 5.8F ($k_{max} = 3$ m^{-1} and $L = 1$ m). In general, I expect grid orientations to be uniformly distributed only in infinite-sized environments, or in environments that are much larger than the pattern size. Nevertheless, because grid orientation depends on the boundary conditions, it remains difficult to compare the distributions obtained here with the ones observed experimentally (Hafting et al., 2005; Stensola et al., 2012; Krupic et al., 2015; Stensola et al., 2015). Finally, in order to explain grid alignment across cells and/or environments (Hafting et al., 2005; Stensola

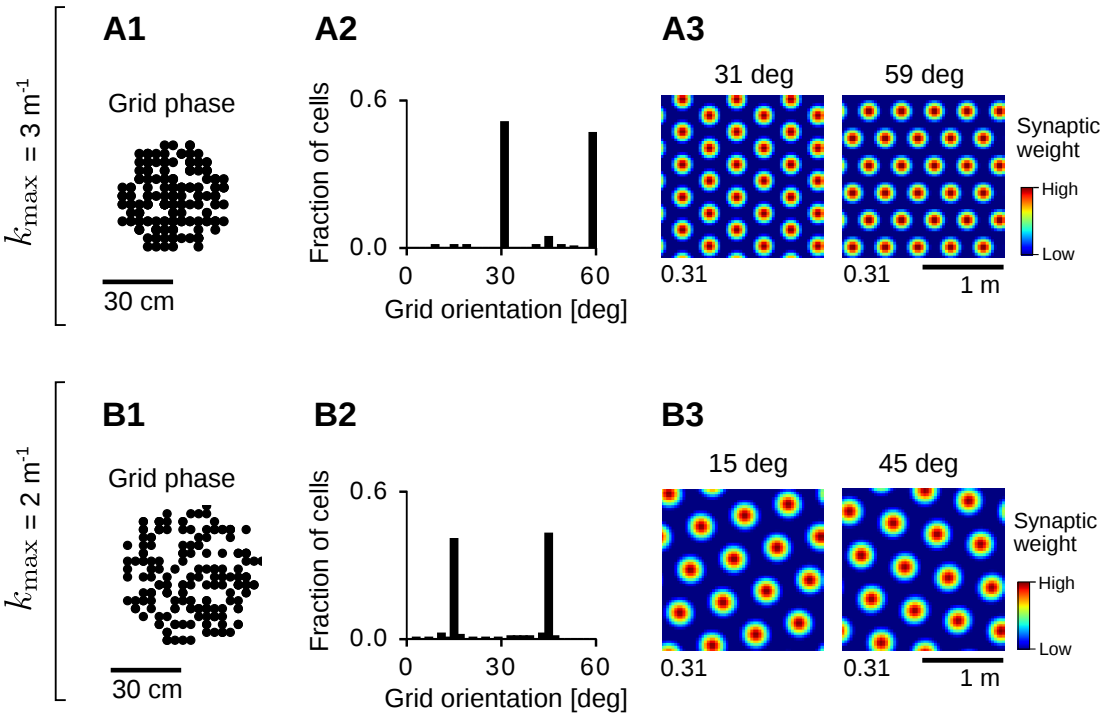

Figure 5.7. | Geometric properties of the grid patterns. A) Distribution of grid spatial phases (A1) and grid orientations (A2) for patterns at frequency $k_{max} = 3$ m^{-1} in an arena of side-length $L = 2$ m ($\sigma = 6.25$ cm, $\tau_L = 0.16$ s; see also Figure 5.6, bottom-left panel). Distributions were obtained from the average weight dynamics in Equation 5.16 for 200 random initial configurations of the synaptic weights ($t = 10^6$ s). Only patterns with gridness scores larger than 0.5 were considered (197/200). Panel A3 shows example weight patterns for the two most common orientations in A2 (maximal values at the bottom-left corner). **B)** Same as in A but for patterns at spatial frequency $k_{max} = 2$ m^{-1} in an arena of side-length $L = 2$ m ($\sigma = 6.25$ cm, $\tau_L = 0.35$ s; see also Figure 5.6, bottom-right panel). A fraction of 182/200 grids had a gridness score larger than 0.5. See Section 5.A.6 for further details and parameter values.

et al., 2012), collateral interactions between developing grid cells may be required (Si et al., 2012; Si and Treves, 2013; Urdapilleta et al., 2017, Sections 5.5.4 and 4.3.1).

5.4.3. Pattern formation with spatially-irregular inputs

In Section 5.4.1, I demonstrated the emergence of grid-like patterns in the case of *spatially-regular* inputs, i.e., for each input cell having a single Gaussian receptive field in space (Section 5.2.3). I now show that similar results are obtained in the case of *spatially-irregular* inputs (Section 5.2.3). I generate spatially-irregular input tuning curves Ψ_i^{in} by superimposing $M > 1$ Gaussian functions with equal width σ, but random centers and random amplitudes (Equation 5.10, see Figure 5.8A for examples). The functions Ψ_i^{in} are normalized such that their average firing rate r_{av} is constant for all input neurons and independent from the number M of superimposed receptive

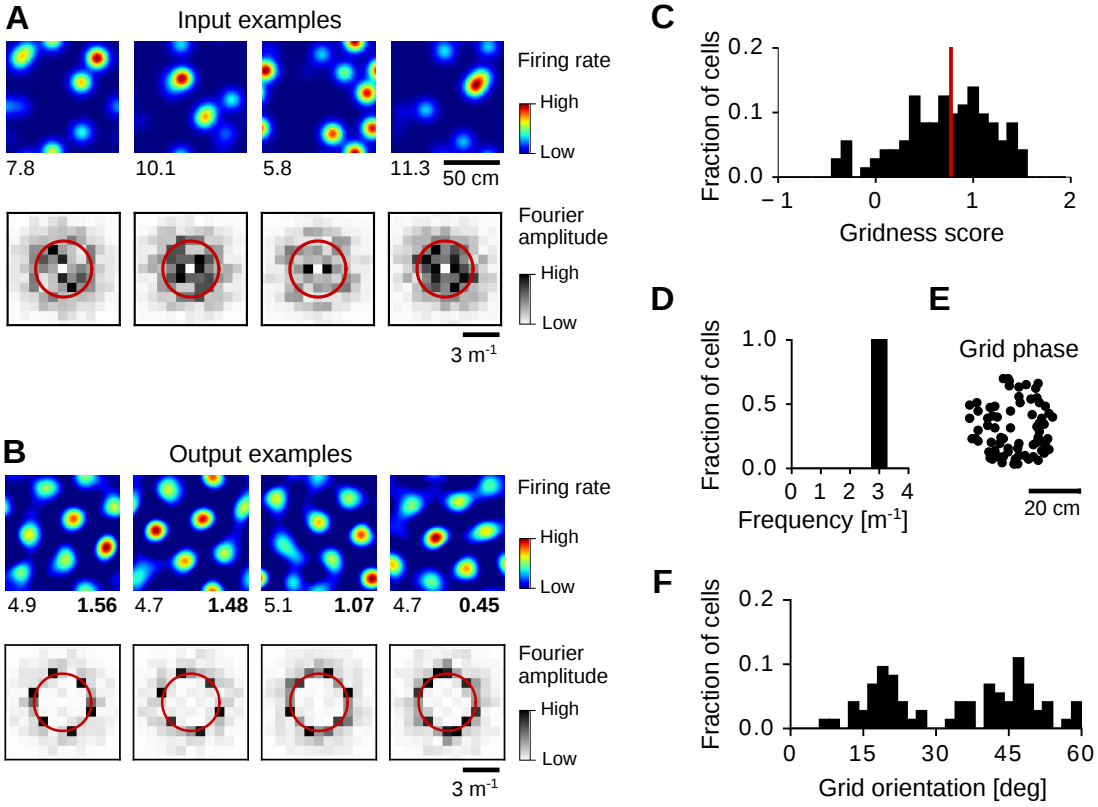

Figure 5.8. | Grid-pattern formation with spatially-irregular inputs. A) Four examples of irregular input firing-rate maps (top row) and the corresponding Fourier spectra (bottom row). The maximal firing rate (spikes/s) is reported at the bottom-left corner. The red circles indicate the spatial frequency $k_{\max} = 3\ \mathrm{m}^{-1}$. **B)** Four examples of output firing-rate maps (top row) and the corresponding Fourier spectra (bottom row). The gridness score is reported at the bottom-right corner. Output firing-rate maps were estimated from the average weight dynamics in Equation 5.16 ($t = 10^6$ s) for four different realizations of the spatial inputs. **C-F)** Distribution of gridness scores (C), grid spatial frequencies (D), grid spatial phases (E), and grid orientations (F) for 100 random realizations of the spatial inputs. The red vertical line in C indicates the mean score (0.77). See Section 5.A.6 for further details and parameter values.

fields.

I test grid-pattern formation in this scenario by simulating the average dynamics of the synaptic weights (Equations 5.16 and 5.21) for random realizations of the input tuning curves Ψ_i^{in}, with $N = 3600$ input neurons and $M = 10$ receptive fields per neuron. I then estimate output firing-rate maps from the synaptic weights at the end of the simulations ($t = 10^6$ s). The results are shown in Figure 5.8 B and C. In the majority of the cases (73/100) a regular grid-like pattern emerges at the output.

Like in the case of spatially-regular inputs, the spatial scale of the output patterns depends on the long adaptation time constant τ_{L} and on the width σ of the input receptive fields. Indeed, for $\sigma = 6.25$ cm and $\tau_{\text{L}} = 0.16$ s, I obtain output grid patterns with spatial frequency $k_{\max} = 3$ m^{-1} (Figure 5.8D), which is equal to the one obtained for spatially-regular inputs with the same parameter values (Figure 5.4 and Figure 5.6, bottom-left panel). This can be understood by the fact that the expected eigenvalue spectrum for spatially-irregular inputs $\langle \lambda_{\text{irr}}(k) \rangle$ is qualitatively similar to the eigenvalue spectrum $\lambda(k)$ for spatially-regular inputs (Section 5.A.4, Equation 5.72):

$$\langle \lambda_{\text{irr}}(k) \rangle \approx \Phi \lambda(k) + \text{const.} \tag{5.33}$$

where $0 \leq \Phi \leq 1$ is a scale factor. I also find that the scale factor Φ depends on the number M of superimposed fields, i.e., $\Phi \approx 4/(3M)$ for $M > 3$ (Equation 5.82), meaning that structure formation is slower for larger numbers of superimposed fields (Equation 5.29).

Finally, like in the case of spatially-regular inputs, with periodic boundary conditions the spatial phases of the simulated grids distribute evenly in the arena (Figure 5.8E), and the grid orientations tend to cluster according to the grid scale and the size of the environment (Figure 5.8F, see also Figure 5.7A2 for the same grid scale in a larger environment).

5.5. Discussion

I studied the origin of grid-cell patterns in a single-cell spiking model relying solely on 1) spatially-tuned feed-forward inputs, 2) spike-rate adaptation, 3) and synaptic plasticity at the input synapses. I considered two input scenarios: spatially-regular inputs (reminiscent of place-cell activity), and spatially-irregular inputs (reminiscent of parasubicular activity). First, I studied the average dynamics of the system analytically, and I derived necessary conditions for the emergence of spatially-periodic solutions (Section 5.3). I then simulated the model numerically, and showed that grid-like patterns emerge both with spatially-regular and spatially irregular inputs (Section 5.4). In the following, I discuss the main assumptions and predictions of my model.

5.5.1. Input spatial tuning and the origin of grid-cell patterns

I assumed that the feed-forward input activity is spatially tuned. Such spatial tuning could be provided by hippocampal place cells, or by other cortical or sub-cortical structures with less regular spatial tuning. From a theoretical point of view, I find that grid patterns emerge faster with place-cell-like inputs, i.e., with inputs having a single receptive field in space. From an anatomical point of view, both scenarios seem plausible. On the one hand, grid-cell activity requires excitatory drive from the hippocampus (Bonnevie et al., 2013), which projects to the deep layers of the MEC (Tamamaki and Nojyo, 1995; Sürmeli et al., 2015) where grid cells are found (Sargolini et al., 2006; Boccara et al., 2010). On the other hand, parasubicular inputs target layer II of the MEC (Van Groen and Wyss, 1990; Caballero-Bleda and Witter, 1993, 1994; Canto et al., 2012; Tang et al., 2016) where grid cells are most abundant (Sargolini et al., 2006; Boccara et al., 2010). Although a small fraction of parasubicular cells already shows grid-like tuning (Boccara et al., 2010; Tang et al., 2016), the activity in parasubiculum is often characterized by multiple spatially-irregular fields (Cacucci et al., 2004; Hargreaves et al., 2005, 2007; Boccara et al., 2010; Tang et al., 2016) similar to those assumed in my model (Figure 5.8).

That grid-cell activity could originate from parasubicular inputs is further supported by the detailed layout of the entorhinal circuit. Layer II principal neurons segregate into stellate and pyramidal cells, which are distinguished by their morphology, intrinsic properties (Alonso and Klink, 1993), and immunoreactivity (Varga et al., 2010; Ray et al., 2014; Kitamura et al., 2014). Interestingly, pyramidal-cell somata cluster into anatomical patches (Ray et al., 2014; Kitamura et al., 2014), which are preferentially targeted by parasubicular axons (Tang et al., 2016). And the spiking activity in parasubiculum precedes the activity of layer II pyramidal cells by a few degrees in the theta cycle (Tang et al., 2016). Such a network configuration suggests that grid patterns may originate in the layer II pyramidal cells via parasubicular inputs, and be inherited by the stellate cells via feed-forward projections. Consistent with this view is that both stellate and pyramidal cells show grid spatial tuning (Sun et al., 2015), and that direct intra-laminar connections are found from pyramidal onto stellate cells and not vice-versa (Fuchs et al., 2016; Winterer et al., 2017); but see (Donato et al., 2017).

In summary, my model is consistent with entorhinal grid-cell activity originating either in the superficial layers via parasubicular input or in the deep layers via hippocampal input. It is also possible that multiple sites of origin exist, and that grid-like tuning is inherited—and even sharpened—via feed-forward projections from the deep to the superficial layers (Iijima et al., 1996; Van Haeften et al., 2003; Kloosterman et al., 2003; Beed et al., 2010; Tocker et al., 2015) or from the superficial to the deep layers (Sürmeli et al., 2015). The reader is kindly referred to Chapter 6 for results on the

inheritance and amplification of grid-cell patterns.

5.5.2. Spike-rate adaptation

My model relies on the presence of a spike-rate adaptation mechanism. Spike-rate adaptation has been observed throughout the cortex (La Camera et al., 2006), and is prominent in layer II of the MEC, in both stellate and pyramidal neurons (Alonso and Klink, 1993; Van der Linden and da Silva, 1998, Section 2.2.2). Yoshida et al. (2013) also reported a dorso-ventral gradient in the adaptation strength of layer II entorhinal cells. However, because adaptation was found to be stronger ventrally than dorsally, Yoshida et al. (2013) interpreted their results as evidence against grid-cell models based on adaptation. Yet the critical variable controlling the grid scale is not the strength of adaptation, but rather its temporal dynamics (Figure 2.1), which was not systematically analyzed (Yoshida et al., 2013); see also (Urdapilleta et al., 2017) for a similar discussion on this point.

I modeled spike-rate adaptation by applying a temporal kernel K to the input spike trains (Equation 5.1). The kernel K, was composed of a brief depolarization peak and a slower hyper-polarizing potential (on a time scale of hundreds of milliseconds). Such a slow hyper-polarizing potential reduced the output firing rate in response to persistent excitation, and it filtered the input activity in a low-frequency band (i.e. with a resonance frequency of about 1 Hz, see Figure 5.1B2). The shape of the kernel was motivated by long-lasting hyper-polarizing potentials following excitatory postsynaptic potentials found in hippocampal CA1 pyramidal neurons (Nicoll and Alger, 1981), although similar responses have not been observed in the MEC yet.

However, the formation of grid-cell patterns could rely on any other cellular or synaptic mechanism that effectively acts as a band-pass filter on the input activity. A candidate mechanism is the after-spike hyperpolarizing potential (AHP). AHPs are indeed observed in the superficial layers of the MEC where single action potentials are followed by both a fast (2-5 ms) and a medium AHP (20-100 ms Alonso et al., 1990; Alonso and Klink, 1993; Pastoll et al., 2012, Section 2.2.2). To assess whether such hyperpolarizing potentials could underlie grid-pattern formation, I extended my model to account for AHPs (Section 5.A.5). However, I found that grids at typical spatial scales cannot be obtained by AHPs alone. Yet after-spike potentials could amplify the effects of a band-pass filtering mechanism that is already present at the input.

Spike-rate adaptation could also rely on hyperpolarization-activated cation currents (I_h), which depend on HCN channels (Accili et al., 2002; Robinson and Siegelbaum, 2003). Fast I_h currents (mediated by HCN1 channels) have been shown to control the theta-frequency resonance of entorhinal stellate cells in vitro (Richter et al., 2000; Dickson et al., 2000; Nolan et al., 2007; Giocomo and Hasselmo, 2008b; Pastoll et al.,

2012). Instead, slower I_h currents (mediated by HCN2-4 channels) could generate in entorhinal cells the low-frequency resonance assumed by my model (Figure 5.1B2).

5.5.3. Synaptic plasticity

I propose that grid-cell patterns emerge from a synaptic reorganization of the MEC network, which is assumed to be plastic. This is in line with both LTP and LTD being reported in the entorhinal cortex (Alonso et al., 1990; de Curtis and Llinas, 1993; Yun et al., 2002; Solger et al., 2004; Yang et al., 2004), but see also (Deng and Lei, 2007). Additionally, asymmetric STDP was observed in the MEC (Zhou et al., 2005). Although I used a symmetric learning window, the exact window shape has little effect on grid-pattern formation, provided that its temporal width (on the order of tens of milliseconds) is much shorter than the correlation length of the input activities (on the order of hundreds of milliseconds). However, the window integral shall be strictly positive (Equation 5.32), i.e., an odd learning window would not lead to any learning in this model.

Structure formation via Hebbian learning is typically a slow process. In my model, grid-like patterns emerge on a time scale that is inversely proportional to the learning rate η and to the maximal eigenvalue $\lambda_{\max}$ (Equation 5.29). The latter depends on the spatial density $\rho = N/L^2$ of input receptive fields, on the integral W_{tot} of the learning window, on the shapes of the input-tuning curves $\mathcal{G}$, and on the dynamics of the adaptation kernel K (Equation 5.32). Because most of these quantities are under-constrained by empirical data, a direct comparison with experimental time scales remains difficult. Yet learning shall be slow enough such that the input correlations that drive structure formation dominate over random fluctuations of the synaptic weights, which are due to the random walk of the virtual rat and the shot noise of the stochastic spiking. In my simulations, I find that this requires tens of hours of spatial exploration (Figure 5.5A).

Such slow process may seem in contrast with grid-cell activity appearing immediately in a novel environment (Hafting et al., 2005; Fyhn et al., 2007). However, grid-like tuning may not need to be learned in each environment anew, but rather recalled—and possibly refined—from the experience of similar environments explored in the past. Although hippocampal place cells (Muller and Kubie, 1987; Bostock et al., 1991) and entorhinal non-grid spatial cells (Diehl et al., 2017) seem to remap completely in novel spaces, less is known on the remapping of parasubicular cells. Additionally, grid-cell learning could generalize across spatial contexts through border and boundary-vector inputs (Solstad et al., 2008; Lever et al., 2009), which are invariant across environments.

I suggest that a structure in the synaptic weights may be formed during the animal's ontogenetic development, i.e., within a two-week period after the animal leaves the

nest (Langston et al., 2010; Wills et al., 2010, 2012). Consistent with this hypothesis is that stable spatial firing is observed before grid-cell maturation, e.g., hippocampal place cells develop prior to grid cells (Langston et al., 2010; Wills et al., 2010), and that grid patterns are disrupted in adult animals following knock-out of NMDA receptors in pups (Dagslott et al., 2016; Gil et al., 2018).

5.5.4. Recurrent dynamics

I studied the emergence of grid patterns in a purely single-cell model, ignoring any network-level interaction between the neurons. However, because excitatory and inhibitory recurrent circuits have been described in the MEC (Dhillon and Jones, 2000; Couey et al., 2013; Pastoll et al., 2013; Fuchs et al., 2016; Winterer et al., 2017), grid cells are likely to be mutually coupled (Yoon et al., 2013; Dunn et al., 2015). Such recurrent connections could explain the modular organization of grid-cell properties (Stensola et al., 2012; Urdapilleta et al., 2017, Section 3.1.2) and the coherent remapping across environments (Fyhn et al., 2007; Yoon et al., 2013). Feedback interactions within a module may also amplify an initially broad grid-tuning given by the feed-forward inputs, similarly to the sharpening of receptive fields in visual cortex (Ben-Yishai et al., 1995; Carandini and Ringach, 1997, Chapter 6). Finally, recurrent dynamics may sustain grid-like activity when the feed-forward inputs are temporally untuned, like in attractor models (McNaughton et al., 2006, Section 4.1).

5.5.5. Related models

The present work—and the one by Kropff and Treves (2008)—belong to a broad category of grid-cell models based on spatially-tuned feed-forward inputs and Hebbian synaptic plasticity (Castro and Aguiar, 2014; Stepanyuk, 2015; Dordek et al., 2016; Monsalve-Mercado and Leibold, 2017; Weber and Sprekeler, 2018, Section 4.3.2). In all these models, periodic spatial patterns arise via a common underlying principle: the input correlations that drive the dynamics of the synaptic weights have the form of a Mexican-hat kernel (Figure 5.2). What distinguishes the models among each other—and generates distinct predictions—is the specific mechanism by which such Mexican-hat interactions are obtained.

In my model, a Mexican-hat kernel results from the intrinsic adaptation dynamics of the output neuron, which controls the grid scale directly (Figure 5.6 and Section 5.5.6).

By contrast, in the models by Castro and Aguiar (2014) and Stepanyuk (2015), Mexican-hat correlations arise from the learning rule itself, i.e., by assuming that synaptic plasticity switches between LTP and LTD based on pre- and postsynaptic activities (Bienenstock et al., 1982). In this case, the grid spatial scale shall be affected by interfering with the learning rule.

In a different model, Dordek et al. (2016) obtain Mexican-hat correlations by constraining the input activity to be effectively zero-mean. The authors discuss that such a zero-mean constraint could originate either from lateral inhibition or from a zero-mean temporal filter controlling the output activity of the neuron. In the latter case, the model by Dordek et al. (2016) is analogous to the present one. I note, however, that effectively zero-mean inputs are neither necessary nor sufficient for grid patterns to emerge. Instead, pattern formation depends on the dynamics of the temporal filter and on the shape of the input tuning curves, but not on their means. This can be easily understood by considering the system's eigenvalue spectrum in Fourier space (Equation 5.30), where the zero-frequency mode ($k = 0$) is not relevant for the emergence of spatially-periodic solutions. Also note that the smallest grid scales in my model are obtained with negative-mean temporal filters (Figure 5.3). Yet my results agree with the ones of Dordek et al. (2016) in that the non-linearity introduced by imposing non-negative synaptic weights is sufficient for a triangular symmetry to emerge.

Alternatively, Mexican-hat correlations could emerge from phase-precessing feed-forward inputs (Monsalve-Mercado and Leibold, 2017). In this case, grid-cell activity shall be impaired when phase precession is disrupted.

Finally, Weber and Sprekeler (2018) proposed a model where the interplay between spatially-narrow feed-forward excitation and spatially-broad feed-forward inhibition generates a Mexican-hat kernel. This model predicts that the grid scale shall be affected by manipulating inhibitory inputs to the MEC.

5.5.6. Model predictions and conclusion

I presented a single-cell model for the origin of grid-cell activity based on Hebbian synaptic plasticity and spike-rate adaptation. My work builds upon the model by Kropff and Treves (2008) and improves its original formulation in several aspects: 1) grid-like patterns emerge form a purely single-cell mechanism independently of any network-level interaction; 2) neuronal activities are spike-based and stochastic; 3) the input synaptic weights are purely excitatory; 4) the dynamics of the synaptic weights is studied analytically and linked to classical Turing-like patterns.

The present model makes similar experimental predictions as the model by Kropff and Treves (2008). First, grid-cell patterns shall be affected by disrupting synaptic plasticity during ontogenesis. Consistently, abolishing NMDA receptors in young animals results in degraded grid patterns in adulthood Dagslott et al. (2016); Gil et al. (2018). Second, grid-cell activity shall be influenced by systematic behavioral or environmental biases in the first weeks of spatial exploration, e.g., by rising animals in environments without boundaries or with non-zero surface curvature (Kruge et al., 2016; Stella et al., 2013). Third, the grid scale shall be affected by three factors: 1) the

average speed of the rat during ontogenetic development; 2) the time constant of the recovery from spike-rate adaptation; 3) the spatial tuning-width of the feed-forward inputs. Fourth, grids at larger scales shall develop faster as compared to grids at smaller scales (Figure 5.3). Note that the dependence of the grid scale on the input-tuning width and the relative time of spatial pattern formation have not been explicitly studied in previous models, and they thus constitute novel predictions.

I believe that manipulations of the intrinsic adaptation properties of single cells are key to distinguish my model from other feed-forward models based on Hebbian learning (Section 5.5.5). To this end, further experimental work shall be devoted to pinpoint the biophysical mechanisms underlying adaptation in the MEC. Extensions of the present model could also explain how the geometry of the enclosure affects grid-cell symmetry (Krupic et al., 2015, 2018; Wernle et al., 2018), and how grid-like tuning emerges in non-spatial contexts (Constantinescu et al., 2016; Aronov et al., 2017).

To conclude, this study contributes to a better understanding of the fundamental principles governing grid-cell activity, and lays the groundwork for more biophysically-realistic grid-cell models.

5.6. Chapter Summary

In this chapter, I studied a computational model for the origin of grid-cell patterns. Building upon previous theoretical work, I proposed that periodic representations of space could arise from a single-cell process based on spatially-selective inputs and spike-rate adaptation. Compared to related proposals, my model achieved a higher level of biological realism, gave unprecedented analytical insights, and generated novel predictions. In the next chapter, I will investigate how grid-cell activity may be affected by excitatory microcircuits in the MEC.

5.A. Appendix

5.A.1. Weight normalization

Here, I derive the dynamics of the mean synaptic weight $w_{\text{av}} = N^{-1} \sum_{i=1}^{N} \bar{w}_i$ for a neuron with N synapses and temporally-averaged weights $\bar{w}_i$. I recall the weight dynamics in Equation 5.16

$$\eta^{-1} \frac{\mathrm{d}}{\mathrm{d}t} \bar{w}_i = \sum_{j=1}^{N} C_{ij} \bar{w}_j - a \bar{w}_i + b \quad \text{with} \quad \bar{w}_i \geq 0 \,. \tag{5.34}$$

By taking the average over the index i at both sides of Equation 5.34, I obtain

$$\eta^{-1}\frac{\mathrm{d}}{\mathrm{d}t}w_{\mathrm{av}} = (NC_{\mathrm{av}} - a)w_{\mathrm{av}} + b \tag{5.35}$$

where I defined the mean correlation $C_{\mathrm{av}} := N^{-2}\sum_{ij} C_{ij}$. Note that I used the property $\sum_j C_{ij} = NC_{\mathrm{av}}$ for all i, which holds true for translation-invariant inputs. Therefore, for $NC_{\mathrm{av}} < a$, the mean weight w_{av} decays exponentially with time constant

$$\tau_{\mathrm{av}} := \frac{1}{\eta(a - NC_{\mathrm{av}})} \tag{5.36}$$

to the normalization level

$$w_{\mathrm{av}}^{\infty} := \frac{b}{a - NC_{\mathrm{av}}}\,. \tag{5.37}$$

5.A.2. Input correlation for general inputs

In this section, I estimate the input correlation matrix

$$C_{ij} \overset{(5.20)}{\approx} W_{\mathrm{tot}} \int_0^{\infty} \mathrm{d}\tau\; K(\tau)\overline{r_i^{\mathrm{in}}(t)r_j^{\mathrm{in}}(t-\tau)} \quad \text{with} \quad i,j = 1,\ldots,N \tag{5.38}$$

for general spatial tuning curves Ψ_i^{in} and smooth movement trajectories of the virtual rat (Section 5.2.4). I start by computing the temporal average $\overline{r_i^{\mathrm{in}}(t)r_j^{\mathrm{in}}(t-\tau)}$ of the product between the input activities $r_i^{\mathrm{in}}(t)$ and the delayed input activities $r_j^{\mathrm{in}}(t-\tau)$. I assume that the stochastic process $\mathbf{X}_t$ controlling the virtual-rat trajectory (Equation 5.11) is ergodic with respect to the auto-covariance, i.e.,

$$\frac{1}{T}\int_0^T \mathbf{d}t\; \mathbf{x}_t\,\mathbf{x}_{t-\tau} = \langle \mathbf{X}_t, \mathbf{X}_{t-\tau}\rangle \quad \text{for} \quad T \to \infty \tag{5.39}$$

where the angular brackets denote statistical expectation. By using this ergodicity property (Equation 5.39) and the spatial tuning of the inputs (Equation 5.7), I derive

$$\overline{r_i^{\mathrm{in}}(t)r_j^{\mathrm{in}}(t-\tau)} = \overline{\Psi_i^{\mathrm{in}}(\mathbf{x}_t)\Psi_j^{\mathrm{in}}(\mathbf{x}_{t-\tau})} \approx \langle \Psi_i^{\mathrm{in}}(\mathbf{X}_t)\Psi_j^{\mathrm{in}}(\mathbf{X}_{t-\tau})\rangle\,. \tag{5.40}$$

Note that Equation 5.40 is only valid in an approximate sense because Equation 5.39 assumes $T \to \infty$, but the averaging time window has finite length $T \ll \tau_{\mathrm{str}}$ where τ_{str}

is the structure-formation time constant (Equation 5.29). From Equation 5.40 follows

$$\begin{aligned}
\overline{r_i^{\text{in}}(t) r_j^{\text{in}}(t-\tau)} &\approx \left\langle \Psi_i^{\text{in}}(\mathbf{X}_t) \Psi_j^{\text{in}}(\mathbf{X}_{t-\tau}) \right\rangle && (5.41)\\
&:= \iint \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{x}'\, \Psi_i^{\text{in}}(\mathbf{x}) \Psi_j^{\text{in}}(\mathbf{x}')\, p(\mathbf{x}, t, \mathbf{x}', t-\tau) && (5.42)\\
&= \iint \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{x}'\, \Psi_i^{\text{in}}(\mathbf{x}) \Psi_j^{\text{in}}(\mathbf{x}')\, p(\mathbf{x}', t-\tau|\mathbf{x}, t) p(\mathbf{x}, t) && (5.43)\\
&= \frac{1}{L^2} \iint \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{x}'\, \Psi_i^{\text{in}}(\mathbf{x}) \Psi_j^{\text{in}}(\mathbf{x}')\, p(\mathbf{x}', t-\tau|\mathbf{x}, t) && (5.44)
\end{aligned}$$

where the integrals in Equations 5.42-5.44 run over all positions in the environment (a square arena of side-length L), and $p(\mathbf{x}, t, \mathbf{x}', t-\tau)$ is the joint probability density of the virtual rat being at position $\mathbf{x}$ at time t and at position $\mathbf{x}'$ at time $t-\tau$. From Equation 5.43 to Equation 5.44, I used the fact that, for large times t, the virtual rat has equal probability of being in any position $\mathbf{x}$, i.e., $p(\mathbf{x}, t) = 1/L^2$.

Equation 5.44 shows that the temporal average $\overline{r_i^{\text{in}}(t) r_j^{\text{in}}(t-\tau)}$ can be estimated from the input tuning curves Ψ_i^{in} and Ψ_j^{in}, and the conditional probability density $p(\mathbf{x}', t-\tau|\mathbf{x}, t)$. This conditional probability density has not yet been solved for correlated random walks in two dimensions (Codling et al., 2008). Nevertheless, an additional approximation is possible. Because the temporal average $\overline{r_i^{\text{in}}(t) r_j^{\text{in}}(t-\tau)}$ is weighted by the adaptation adaptation kernel $K(\tau)$ (Equation 5.38), and $K(\tau)$ is negligible for $\tau > \tau_{\max} \approx 5\tau_{\text{L}}$ (Equation 5.2), I am interested in the conditional probability $p(\mathbf{x}', t-\tau|\mathbf{x}, t)$ only at lags $\tau < \tau_{\max}$. In this case, for movement trajectories that are sufficiently smooth, I can assume that in a time τ the virtual rat has moved to a position $\mathbf{x}$ at distance $|\mathbf{x} - \mathbf{x}'| = \tau v$ from the initial position $\mathbf{x}'$, that is

$$p(\mathbf{x}', t-\tau|\mathbf{x}, t) \approx \frac{\delta(|\mathbf{x} - \mathbf{x}'| - \tau v)}{2\pi\tau v} \tag{5.45}$$

where v is the speed of the virtual rat (Equation 5.11), and the denominator ensures that $\int \mathrm{d}\mathbf{x}' p(\mathbf{x}', t-\tau|\mathbf{x}, t) = 1$; see also Figure 5.1D for exemplary virtual-rat trajectories in this scenario. I now use Equation 5.45 in Equation 5.44, and let $\mathbf{z} := \mathbf{x}' - \mathbf{x}$:

$$\overline{r_i^{\text{in}}(t) r_j^{\text{in}}(t-\tau)} \approx \frac{1}{L^2} \underbrace{\int \mathrm{d}\mathbf{z}\, \frac{\delta(|\mathbf{z}| - \tau v)}{2\pi\tau v}}_{=: \oint_{|\mathbf{z}|=\tau v} \mathrm{d}\mathbf{z}} \underbrace{\int \mathrm{d}\mathbf{x}\; \Psi_i^{\text{in}}(\mathbf{x}) \Psi_j^{\text{in}}(\mathbf{x}+\mathbf{z})}_{=: \Psi_i^{\text{in}} \star \Psi_j^{\text{in}}\big|_{\mathbf{z}}} . \tag{5.46}$$

From Equation 5.46, the temporal average $\overline{r_i^{\text{in}}(t) r_j^{\text{in}}(t-\tau)}$ is approximated by the integral of the spatial cross-correlation $\Psi_i^{\text{in}} \star \Psi_j^{\text{in}}$ over a circle of radius τv. Finally, by

using Equation 5.46 in Equation 5.20, I obtain

$$C_{ij} \approx \frac{W_{\text{tot}}}{L^2} \int_0^\infty \mathrm{d}\tau\, K(\tau) \oint_{|\mathbf{z}|=\tau v} \mathrm{d}\mathbf{z}\; \Psi_i^{\text{in}} \star \Psi_j^{\text{in}} \Big|_{\mathbf{z}} . \tag{5.47}$$

5.A.3. Input correlation for spatially-regular inputs

In this section, I compute the input correlation function C and its Fourier spectrum $\widehat{C}$ in the case of spatially-regular inputs (see Section 5.3.2). First, I rewrite the input correlation matrix C_{ij} in Equation 5.21 as a continuous function $C(\mathbf{r}, \mathbf{r}')$ by labeling neurons according to their receptive-field centers $\mathbf{r}$ and $\mathbf{r}'$:

$$C(\mathbf{r}, \mathbf{r}') \approx \frac{W_{\text{tot}}}{L^2} \int_0^\infty \mathrm{d}\tau\, K(\tau) \oint_{|\mathbf{z}|=\tau v} \mathrm{d}\mathbf{z}\; \Psi_{\mathbf{r}}^{\text{in}} \star \Psi_{\mathbf{r}'}^{\text{in}} \Big|_{\mathbf{z}} \tag{5.48}$$

where $\Psi_{\mathbf{r}}^{\text{in}}(\mathbf{x}) := \mathcal{G}(|\mathbf{x} - \mathbf{r}|)$ is a Gaussian input tuning curve centered at position $\mathbf{r}$ (Equation 5.9). Because the inputs are translation invariant, the correlation function C depends only on the translation vector $\mathbf{u} := \mathbf{r} - \mathbf{r}'$:

$$C(\mathbf{r}, \mathbf{r}') = C(\mathbf{u}, \mathbf{0}) = C(\mathbf{u}) \;\approx\; \frac{W_{\text{tot}}}{L^2} \int_0^\infty \mathrm{d}\tau\, K(\tau) \oint_{|\mathbf{z}|=\tau v} \mathrm{d}\mathbf{z}\; \Psi_{\mathbf{u}}^{\text{in}} \star \Psi_{\mathbf{0}}^{\text{in}} \Big|_{\mathbf{z}} \tag{5.49}$$

$$= \frac{W_{\text{tot}}}{L^2} \int_0^\infty \mathrm{d}\tau\, K(\tau) \oint_{|\mathbf{z}|=\tau v} \mathrm{d}\mathbf{z}\; \Psi_{\mathbf{0}}^{\text{in}} \star \Psi_{\mathbf{0}}^{\text{in}} \Big|_{\mathbf{u}+\mathbf{z}} \tag{5.50}$$

where $\Psi_{\mathbf{0}}^{\text{in}}(\mathbf{x}) := \mathcal{G}(|\mathbf{x}|)$ is the tuning curve centered at the origin $\mathbf{0} = (0, 0)$. Next, I substitute in Equation 5.50 the definition of the integral operator in Equation 5.46:

$$C(\mathbf{u}) \;\approx\; \frac{W_{\text{tot}}}{L^2} \int_0^\infty \mathrm{d}\tau\, K(\tau) \int \mathrm{d}\mathbf{z}\, \frac{\delta(|\mathbf{z}| - \tau v)}{2\pi\tau v}\; \Psi_{\mathbf{0}}^{\text{in}} \star \Psi_{\mathbf{0}}^{\text{in}} \Big|_{\mathbf{u}+\mathbf{z}} . \tag{5.51}$$

It is easy to see that the auto-correlation of a Gaussian is still a Gaussian:

$$\Psi_{\mathbf{0}}^{\text{in}} \star \Psi_{\mathbf{0}}^{\text{in}} \Big|_{\mathbf{u}} = \frac{L^4 r_{\text{av}}^2}{4\pi\sigma^2} \exp\left(-\frac{|\mathbf{u}|^2}{4\sigma^2}\right) \tag{5.52}$$

from which I derive

$$\Psi_{\mathbf{0}}^{\text{in}} \star \Psi_{\mathbf{0}}^{\text{in}} \Big|_{\mathbf{u}+\mathbf{z}} = \frac{L^4 r_{\text{av}}^2}{4\pi\sigma^2} \left[\exp\left(-\frac{|\mathbf{u}|^2 + |\mathbf{z}|^2}{4\sigma^2}\right) \exp\left(-\frac{|\mathbf{u}||\mathbf{z}|\cos(\varphi)}{2\sigma^2}\right)\right] \tag{5.53}$$

where φ is the angle between the vectors $\mathbf{u}$ and $\mathbf{z}$. Finally, by expressing in polar coordinates the vector $\mathbf{z} := |\mathbf{z}|[\cos(\varphi), \sin(\varphi)]$, from Equations 5.51 and 5.53, I obtain

$$C(\mathbf{u}) \approx \frac{W_{\text{tot}} L^2 r_{\text{av}}^2}{4\pi\sigma^2} \int_0^\infty \mathrm{d}\tau\, K(\tau) \exp\left(-\frac{|\mathbf{u}|^2 + (\tau v)^2}{4\sigma^2}\right) I_0\left(-\frac{|\mathbf{u}|\tau v}{2\sigma^2}\right) \tag{5.54}$$

where $I_0(x) := 1/(2\pi) \int_0^{2\pi} \mathrm{d}\varphi \exp(x\cos(\varphi))$ is the zeroth-order modified Bessel function of the first kind.

Fourier spectrum of the input correlation function

Here, I compute the Fourier spectrum of the correlation function C in Equation 5.51. First, I observe that the second integral in Equation 5.51 is a two-dimensional cross-correlation in the variable $\mathbf{z}$ between the functions $\delta(|\mathbf{z}| - \tau v)$ and $\Psi_{\mathbf{0}}^{\mathrm{in}} \star \Psi_{\mathbf{0}}^{\mathrm{in}}\Big|_{\mathbf{z}}$ evaluated at point $\mathbf{u}$. Therefore, by taking the two-dimensional Fourier transform with respect to $\mathbf{u}$ at both sides of Equation 5.51 yields

$$\widehat{C}(\mathbf{k}) \approx \frac{W_{\mathrm{tot}}}{L^2} |\widehat{\Psi}_{\mathbf{0}}^{\mathrm{in}}(\mathbf{k})|^2 \int_0^\infty \mathrm{d}\tau\, K(\tau) \underbrace{\int \mathrm{d}\mathbf{z}\, \frac{\delta(|\mathbf{z}| - \tau v)}{2\pi\tau v} \exp(2\pi j\, \mathbf{z}\cdot\mathbf{k})}_{= J_0(\tau v|\mathbf{k}|)} \tag{5.55}$$

where I defined the Fourier transform pair:

$$\widehat{C}(\mathbf{k}) := \int \mathrm{d}\mathbf{u}\, C(\mathbf{u}) \exp(-2\pi j\, \mathbf{k}\cdot\mathbf{u}) \quad ; \quad C(\mathbf{u}) = \frac{1}{(2\pi)^2} \int \mathrm{d}\mathbf{k}\, \widehat{C}(\mathbf{k}) \exp(2\pi j\, \mathbf{k}\cdot\mathbf{u}) \tag{5.56}$$

with $\mathbf{k}\cdot\mathbf{u} = |\mathbf{k}||\mathbf{u}|\cos(\theta)$, and I used the definition of the zeroth-order Bessel function

$$J_0(k) := \frac{1}{2\pi} \int_0^{2\pi} \mathrm{d}\theta \exp(2\pi j k \cos(\theta)) \,. \tag{5.57}$$

Because the tuning function $\Psi_{\mathbf{0}}^{\mathrm{in}}(\mathbf{x}) := \mathcal{G}(|\mathbf{x}|)$ is circularly symmetric, its two-dimensional Fourier transform $\widehat{\Psi}_{\mathbf{0}}^{\mathrm{in}}(\mathbf{k})$ is proportional to the zeroth-order Hankel transform of $\mathcal{G}$:

$$\widehat{\Psi}_{\mathbf{0}}^{\mathrm{in}}(\mathbf{k}) = 2\pi\, \tilde{\mathcal{G}}(k) \quad \text{with} \quad k := |\mathbf{k}| \,, \tag{5.58}$$

where I defined the zeroth-order Hankel transform pair:

$$\tilde{\mathcal{G}}(k) := \int_0^\infty \mathrm{d}r\, r\, \mathcal{G}(r) J_0(kr) \quad \text{and} \quad \mathcal{G}(r) = \int_0^\infty \mathrm{d}k\, k\, \tilde{\mathcal{G}}(k) J_0(kr) \,. \tag{5.59}$$

By using Equation 5.58 in Equation 5.55, I obtain

$$\widehat{C}(\mathbf{k}) = W_{\mathrm{tot}} \frac{4\pi^2}{L^2} \tilde{\mathcal{G}}^2(k) \int_0^\infty \mathrm{d}\tau\, K(\tau)\, J_0(\tau v k) \tag{5.60}$$

and by defining the equivalent adaptation kernel in space

$$K_{\mathrm{sp}}(r) := \frac{1}{rv} K\left(\frac{r}{v}\right) \tag{5.61}$$

I find

$$\widehat{C}(\mathbf{k}) = W_{\text{tot}} \frac{4\pi^2}{L^2} \tilde{\mathcal{G}}^2(k) \tilde{K}_{\text{sp}}(k)\,. \tag{5.62}$$

Finally, the zeroth-order Hankel transforms of the Gaussian tuning curve $\mathcal{G}$ (Equation 5.9) and of the adaptation kernel in space K_{sp} (Equations 5.61 and 5.2) read

$$\tilde{\mathcal{G}}(k) = \frac{L^2 r_{\text{av}}}{2\pi} \exp\left(-\frac{k^2\sigma^2}{2}\right) \tag{5.63}$$

$$\tilde{K}_{\text{sp}}(k) = \frac{1}{\tau_{\text{S}} v}\left[k^2 + (\tau_{\text{S}} v)^{-2}\right]^{-0.5} - \frac{\mu}{\tau_{\text{L}} v}\left[k^2 + (\tau_{\text{L}} v)^{-2}\right]^{-0.5}. \tag{5.64}$$

5.A.4. Eigenvalue spectrum for spatially-irregular inputs

In this section, I estimate the expected eigenvalue spectrum $\langle \lambda_{\text{irr}}(\mathbf{k}) \rangle$ for spatially-irregular inputs (Sections 5.2.3 and 5.4.3). I recall that, for spatially-regular inputs, in Section 5.3 I obtained (Equation 5.32):

$$\lambda(\mathbf{k}) \approx \rho \frac{W_{\text{tot}}}{L^2} \underbrace{4\pi^2 \tilde{\mathcal{G}}^2(k)}_{\overset{(5.58)}{=} |\widehat{\Psi}^{\text{in}}_{\mathbf{0}}(\mathbf{k})|^2} \tilde{K}_{\text{sp}}(k) - a \quad \text{with} \quad k := |\mathbf{k}| \neq 0 \tag{5.65}$$

where $\tilde{\mathcal{G}}$ and $\tilde{K}_{\text{sp}}$ are the zeroth-order Hankel transforms of the input tuning curve $\mathcal{G}$ (Equation 5.9) and of the equivalent adaptation kernel in space K_{sp} (Equations 5.31 and 5.61). Note that the parameters ρ, L, W_{tot}, and a do not depend on k. From Equation 5.65, the eigenvalue spectrum $\lambda(\mathbf{k})$ is linearly-related to the input power spectrum $|\widehat{\Psi}^{\text{in}}_{\mathbf{0}}(\mathbf{k})|^2$ where $\Psi^{\text{in}}_{\mathbf{0}}(\mathbf{x}) := \mathcal{G}(|\mathbf{x}|)$ is an input tuning curve centered at the origin $\mathbf{0} := (0,0)$ (Section 5.A.3).

Here, in analogy to Equation 5.65, I assume that the expected eigenvalue spectrum $\langle \lambda_{\text{irr}}(\mathbf{k}) \rangle$ for spatially-irregular inputs is linearly-related to the expected input power $\left\langle |\widehat{\Psi}^{\text{in}}_p(\mathbf{k})|^2 \right\rangle$, that is,

$$\langle \lambda_{\text{irr}}(\mathbf{k}) \rangle \approx \rho \frac{W_{\text{tot}}}{L^2} \left\langle |\widehat{\Psi}^{\text{in}}_p(\mathbf{k})|^2 \right\rangle \tilde{K}_{\text{sp}}(k) - a \quad \text{with} \quad k \neq 0 \tag{5.66}$$

where $\widehat{\Psi}^{\text{in}}_p(\mathbf{k})$ is the two-dimensional Fourier transform of the spatially-irregular tuning curve $\Psi^{\text{in}}_p(\mathbf{x})$, and the angular brackets denote statistical expectation across input realizations (see Equation 5.56 for a definition of the two-dimensional Fourier transform). The validity of this assumption is confirmed numerically at the end of this section.

Let me compute the expected input power spectrum $\left\langle |\widehat{\Psi}^{\text{in}}_p(\mathbf{k})|^2 \right\rangle$. I recall that the input maps $\Psi^{\text{in}}_p(\mathbf{x})$ are obtained by the superimposing M Gaussian receptive fields

(Equation 5.10)

$$\Psi_p^{\text{in}}(\mathbf{x}) := \frac{1}{\beta_p} \sum_{m=1}^{M} A_{pm}\, \mathcal{G}(|\mathbf{x} - \mathbf{r}_{pm}|) \quad \text{for} \quad p = 1, 2, \ldots, N \tag{5.67}$$

with

$$\mathcal{G}(r) \overset{(5.9)}{=} \frac{L^2 r_{\text{av}}}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) \quad \text{and} \quad \beta_p := \sum_{m=1}^{M} A_{pm}\,. \tag{5.68}$$

The field amplitudes $A_{pm} \geq 0$ are uniformly distributed in the range $(0, 1)$, and the receptive field centers $\mathbf{r}_{pm}$ are uniformly distributed in the environment (see Figure 5.8A for examples). From Equation 5.67, I derive

$$|\widehat{\Psi}_p^{\text{in}}(\mathbf{k})| = \frac{2\pi}{\beta_p} \tilde{\mathcal{G}}(k) \underbrace{\left| \sum_{m=1}^{M} A_{pm} \exp(-2\pi j\, \mathbf{r}_{pm} \cdot \mathbf{k}) \right|}_{=:\, \alpha_p} \tag{5.69}$$

where $\tilde{\mathcal{G}}(k)$ is the zeroth-order Hankel transform of the Gaussian function $\mathcal{G}(r)$. In deriving Equation 5.69, I used the shift property of the Fourier transform and the equivalence between the Fourier and the zeroth-order Hankel transforms for circularly-symmetric functions (Equation 5.58). Finally, from Equation 5.69 I obtain

$$\left\langle |\widehat{\Psi}_p^{\text{in}}(\mathbf{k})|^2 \right\rangle = 4\pi^2 \tilde{\mathcal{G}}^2(k) \Phi \quad \text{with} \quad \Phi := \left\langle \frac{\alpha_p^2}{\beta_p^2} \right\rangle . \tag{5.70}$$

Therefore, for spatially-irregular inputs, the expected power spectrum $\left\langle |\widehat{\Psi}_p^{\text{in}}(\mathbf{k})|^2 \right\rangle$ is proportional to the power spectrum $4\pi^2 \tilde{\mathcal{G}}^2(k)$ of a single Gaussian $\mathcal{G}$ with scale factor $\Phi \geq 0$. Note that for $|\mathbf{k}| = 0$ I obtain $\Phi = 1$ (Equations 5.69 and 5.70), which means that the average rate r_{av} is independent of the number M of input receptive fields and their specific spatial arrangement. Using Equation 5.70 in Equation 5.66 yields

$$\langle \lambda_{\text{irr}}(\mathbf{k}) \rangle \approx \rho \frac{W_{\text{tot}}}{L^2} 4\pi^2 \tilde{\mathcal{G}}^2(k) \tilde{K}_{\text{sp}}(k) \Phi - a \quad \text{with} \quad k \neq 0\,. \tag{5.71}$$

Finally, from Equations 5.65 and 5.71 I find (Equation 5.33)

$$\langle \lambda_{\text{irr}}(\mathbf{k}) \rangle \approx \Phi \lambda(\mathbf{k}) + a(1 - \Phi)\,. \tag{5.72}$$

In the next section I estimate the scale factor Φ for $|\mathbf{k}| > 0$.

Approximation of the scale factor Φ

The scale factor

$$\Phi \overset{(5.70)}{:=} \left\langle \frac{\alpha_p^2}{\beta_p^2} \right\rangle \tag{5.73}$$

is the second moment of the ratio of the random variables

$$\alpha_p \overset{(5.69)}{:=} \left| \sum_{m=1}^{M} A_{pm} \exp(-2\pi j \, \mathbf{r}_{pm} \cdot \mathbf{k}) \right| \quad \text{and} \quad \beta_p \overset{(5.68)}{:=} \sum_{m=1}^{M} A_{pm} \tag{5.74}$$

where the field amplitudes $A_{pm} \geq 0$ are independently and uniformly distributed in the range $(0, 1)$ and the field centers $\mathbf{r}_{pm}$ are independently and uniformly distributed in a square of side-length L.

In general, for two random variables x and y, the first order Taylor expansion of the ratio $f(\mathbf{z}) = x/y$ around the expected value $\boldsymbol{\mu} := (\langle x \rangle, \langle y \rangle)$ is

$$f(\mathbf{z}) = f(\boldsymbol{\mu}) + f_x(\boldsymbol{\mu})\Delta_x + f_y(\boldsymbol{\mu})\Delta_y + o(\Delta_x^2) + o(\Delta_y^2) + o(\Delta_x \Delta_y) \tag{5.75}$$

where $\mathbf{z} := (x, y)$, $\Delta_x := x - \langle x \rangle$, $\Delta_y := y - \langle y \rangle$, and f_x and f_y are the derivatives of f with respect to x and y. Therefore

$$\begin{aligned} \operatorname{Var}\left(\frac{x}{y}\right) = \langle [f(\mathbf{z}) - f(\boldsymbol{\mu})]^2 \rangle = f_x^2(\boldsymbol{\mu}) \operatorname{Var}(x) + f_y^2(\boldsymbol{\mu}) \operatorname{Var}(y) + \\ 2 f_x(\boldsymbol{\mu}) f_y(\boldsymbol{\mu}) \operatorname{Cov}(x, y) + \sum_{k=0}^{4} o(\langle \Delta_x^k \Delta_y^{4-k} \rangle) \, . \end{aligned} \tag{5.76}$$

By neglecting the higher-order joint moments $\sum_{k=0}^{4} o(\langle \Delta_x^k \Delta_y^{4-k} \rangle)$ and substituting $f_x(\boldsymbol{\mu}) = 1/\langle y \rangle$ and $f_y(\boldsymbol{\mu}) = -\langle x \rangle / \langle y \rangle^2$ I obtain

$$\operatorname{Var}\left(\frac{x}{y}\right) \approx \frac{\langle x \rangle^2}{\langle y \rangle^2} \left[\frac{\operatorname{Var}(x)}{\langle x \rangle^2} + \frac{\operatorname{Var}(y)}{\langle y \rangle^2} - 2 \frac{\operatorname{Cov}(x, y)}{\langle x \rangle \langle y \rangle} \right] \tag{5.77}$$

and

$$\left\langle \frac{x^2}{y^2} \right\rangle \approx \frac{\langle x \rangle^2}{\langle y \rangle^2} \left[\frac{\langle x^2 \rangle}{\langle x \rangle^2} + \frac{\langle y^2 \rangle}{\langle y \rangle^2} - 2 \frac{\operatorname{Cov}(x, y)}{\langle x \rangle \langle y \rangle} - 1 \right] . \tag{5.78}$$

In the following, I use Equation 5.78 to approximate the scale factor Φ (Equation 5.73).

I start by giving an intuitive interpretation of the random variables α_p and β_p. Consider a M-steps random walk on the complex plane with random directions $\mathbf{r}_{pm} \cdot \mathbf{k}$ and random step sizes A_{pm}. The coefficients α_p measure the total distance traveled by the random walker, and the coefficients β_p measure the total length of the path (Equation 5.74). Note that the larger the number of steps M, the smaller is the correlation between the distance traveled α_p and the total path length β_p, i.e., $|\operatorname{Cov}(\alpha_p, \beta_p)| \ll 1$

for $M \gg 1$. In this case, I can neglect the covariance term in Equation 5.78, and the factor Φ is approximated by knowing only the first two moments of the distributions of α_p and β_p.

For $|\mathbf{k}| > 1/L$, the random directions $\mathbf{r}_{pm} \cdot \mathbf{k}$ (mod 1) are approximately uniformly distributed in the range $(0, 1)$. Therefore, the traveled distance α_p follows a Rayleigh distribution with density (Beckmann, 1962)

$$f(\alpha_p) = \frac{2\alpha_p}{M\langle A_{pm}^2 \rangle} \exp\left(-\frac{\alpha_p^2}{M\langle A_{pm}^2 \rangle}\right) \tag{5.79}$$

where $\langle A_{pm}^2 \rangle = 1/3$ for A_{pm} uniformly distributed in interval $(0, 1)$. Hence, the first two moments of α_p read

$$\langle \alpha_p \rangle = \sqrt{\frac{M\pi}{12}} \quad \text{and} \quad \langle \alpha_p^2 \rangle = \frac{M}{3}\,. \tag{5.80}$$

The total path length β_p is the sum of M random variables uniformly distributed in $(0, 1)$, which follows an Irwin-Hall distribution. Therefore, the first two moments of β_p are

$$\langle \beta_p \rangle = \frac{M}{2} \quad \text{and} \quad \langle \beta_p^2 \rangle = \frac{M + 3M^2}{12}\,. \tag{5.81}$$

Finally, by using Equations 5.80 and 5.81 in Equation 5.73, I obtain

$$\Phi(M) \approx \frac{\pi}{3M}\left(\frac{4}{\pi} + \frac{1}{3M}\right) \quad \text{for} \quad M > 1 \quad \text{and} \quad |\mathbf{k}| > 1/L \tag{5.82}$$

$$\approx \frac{4}{3M} \quad \text{for} \quad M > 3 \quad \text{and} \quad |\mathbf{k}| > 1/L\,. \tag{5.83}$$

Figure 5.9A shows the scale factor Φ as a function of the number M of superimposed Gaussian fields (Equation 5.82). Note that the approximation is more accurate for large values of M, which correspond to lower values of $|\operatorname{Cov}(\alpha_p, \beta_p)|$. Figure 5.9B shows the largest eigenvalue in the system as a function of M. The good match between the theoretical curve and the numerical estimations supports the validity of Equation 5.66. Additionally, Equation 5.66 predicts that, irrespectively of the value of M, the largest eigenvalue $\lambda_{\max} = \lambda(k_{\max})$ is always at the critical frequency of $k_{\max} = 3\ \mathrm{m}^{-1}$ for $\sigma = 6.25$ cm and $\tau_{\mathrm{L}} = 0.16$ s, which matches the numerical results in Figure 5.8.

5.A.5. Pattern formation with after-spike potentials

Here, I study whether grid-like patterns could emerge by means of after-spike hyperpolarizing potentials (see discussion in Section 5.5.2). To this end, I consider a model of the output neural activity that is alternative to the one presented in the main text (Sec-

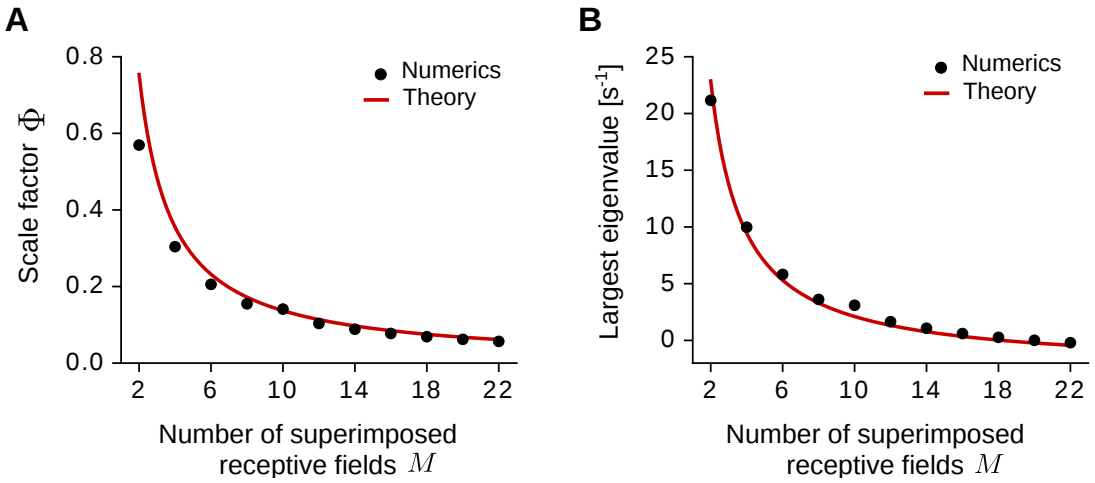

Figure 5.9. | Scale factor Φ and largest eigenvalue $\lambda_{\max}$ for spatially-irregular inputs. **A)** The scale factor Φ for $M > 1$ superimposed fields (Equation 5.70). The black dots are obtained by estimating the power spectrum at frequency $|\mathbf{k}| = 1$ m^{-1} for 3600 input realizations. The red line is the theoretical curve in Equation 5.82. **B)** The largest eigenvalue $\lambda_{\max}$ as a function of the number of superimposed fields M. The black dots are obtained by computing the eigenvalues of the correlation matrix $C_{ij} - a\delta_{ij}$ for $N = 3600$ inputs, where δ_{ij} is the Kronecker delta (Equation 5.21). The red line is obtained from Equations 5.71 and 5.82. Note that, according to Equation 5.71, the largest eigenvalue is always at the critical frequency $k_{\max} = 3$ m^{-1} for any value of M. Parameter values as in Figure 5.8 (see Section 5.A.6).

tion 5.2.1, Equation 5.1). I model input postsynaptic potentials (PSPs) with a kernel K^{in} applied to the input spike trains S_j^{in}, and I model output after-spike hyperpolarizing potentials (AHPs) with a kernel K^{out} applied to the output spike train S^{out}:

$$r^{\text{out}}(t) := r_0 + \int_0^\infty \mathrm{d}s\, K^{\text{out}}(s) S^{\text{out}}(t-s) + \int_0^\infty \mathrm{d}\tau\, K^{\text{in}}(\tau) \sum_{j=1}^N w_j S_j^{\text{in}}(t-\tau) \tag{5.84}$$

where $r_0 \geq 0$ is a baseline firing rate.

First, I show that the average dynamics of Equation 5.84 can be rewritten in terms of an equivalent kernel K^{eq} applied to the input spikes only. I average Equation 5.84 across input and output spike train realizations:

$$\langle r^{\text{out}}(t) \rangle = r_0 + \int_0^\infty \mathrm{d}s\, K^{\text{out}}(s) \langle r^{\text{out}}(t-s) \rangle + \int_0^\infty \mathrm{d}\tau\, K^{\text{in}}(\tau) \sum_{j=1}^N w_j r_j^{\text{in}}(t-\tau)\,. \tag{5.85}$$

By taking the Fourier transform

$$\widehat{f}(\omega) := \int \mathrm{d}t\, f(t) \exp(-j\omega t) \quad ; \quad f(t) = \frac{1}{2\pi} \int \mathrm{d}\omega\, \widehat{f}(\omega) \exp(j\omega t) \tag{5.86}$$

at both sides of Equation 5.85, I obtain

$$\langle \widehat{r}^{\text{out}}(\omega) \rangle = r_0 \delta(\omega) + \frac{\widehat{K}^{\text{in}}(\omega)}{1 - \widehat{K}^{\text{out}}(\omega)} \sum_{j=1}^{N} w_j \widehat{r}_j^{\text{in}}(\omega) \,. \tag{5.87}$$

From Equation 5.85 to Equation 5.87, I assumed that the input and the output kernels are causal, i.e., $K^{\text{in,out}}(t) = 0$ for $t < 0$, and that the output kernel has integral different from 1, i.e., $\widehat{K}^{\text{out}}(0) = \int_0^\infty \mathrm{d}t \, K^{\text{out}}(t) \neq 1$. Finally, by defining the equivalent filter

$$\widehat{K}^{\text{eq}}(\omega) := \frac{\widehat{K}^{\text{in}}(\omega)}{1 - \widehat{K}^{\text{out}}(\omega)} \,, \tag{5.88}$$

the inverse Fourier transform of Equation 5.87 reads

$$\langle r^{\text{out}}(t) \rangle = r_0 + \int_0^\infty \mathrm{d}\tau \, K^{\text{eq}}(\tau) \sum_{j=1}^{N} w_j r_j^{\text{in}}(t - \tau) \,, \tag{5.89}$$

which is equivalent to Equation 5.15 with $K^{\text{eq}} = K$.

Next, I compute the equivalent filter K^{eq} for a simple choice of the input and output kernels

$$K^{\text{in}}(t) := \begin{cases} \dfrac{1}{\tau_{\text{in}}} \exp\left(-\dfrac{t}{\tau_{\text{in}}}\right) & \text{for} \quad t \geq 0 \\ 0 & \text{for} \quad t < 0 \end{cases} \tag{5.90}$$

and

$$K^{\text{out}}(t) := \begin{cases} -\dfrac{\mu_{\text{out}}}{\tau_{\text{out}}} \exp\left(-\dfrac{t}{\tau_{\text{out}}}\right) & \text{for} \quad t \geq 0 \\ 0 & \text{for} \quad t < 0 \end{cases} \tag{5.91}$$

where $\tau_{\text{in}}, \tau_{\text{out}} > 0$ are decay time constants, and the parameter $\mu_{\text{out}} > 0$ scales the integral of the output kernel $\int_0^\infty \mathrm{d}t \, K^{\text{out}}(t) = -\mu_{\text{out}}$. I assume that the input kernel K^{in} (modeling an incoming PSP) decays faster than the output kernel K^{out} (modeling an output AHP), i.e., $\tau_{\text{in}} < \tau_{\text{out}}$. From the definition of the filter K^{eq} in Equation 5.88, I obtain

$$\widehat{K}^{\text{eq}}(\omega) = \underbrace{\frac{1/\tau_{\text{in}}}{1/\tau_{\text{in}} - (1 + \mu_{\text{out}})/\tau_{\text{out}}}}_{=: H} \left[\frac{1/\tau_{\text{in}} - 1/\tau_{\text{out}}}{1/\tau_{\text{in}} + j\omega} - \frac{\mu_{\text{out}}/\tau_{\text{out}}}{(1 + \mu_{\text{out}})/\tau_{\text{out}} + j\omega} \right] \tag{5.92}$$

where I used

$$\widehat{K}^{\text{in}}(\omega) = \frac{1/\tau_{\text{in}}}{1/\tau_{\text{in}} + j\omega} \quad \text{and} \quad \widehat{K}^{\text{out}}(\omega) = -\frac{\mu_{\text{out}}/\tau_{\text{out}}}{1/\tau_{\text{out}} + j\omega} \,. \tag{5.93}$$

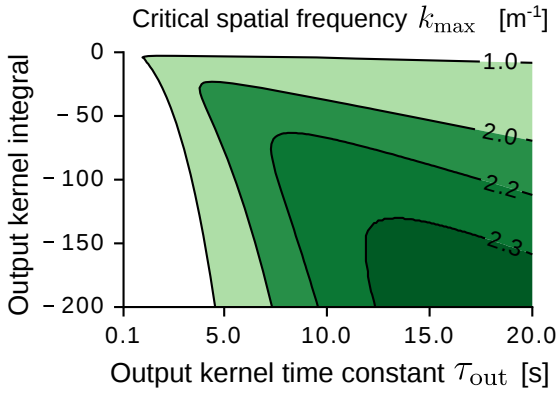

Figure 5.10. | Grid scale with after-spike hyperpolarizing potentials. The critical spatial frequency $k_{\max}$ is plotted as a function of the output-kernel integral $-\mu_{\text{out}}$ and the output-kernel time constant τ_{out} (Equations 5.31 and 5.32 with $K = K^{\text{eq}}$). The black lines are iso-levels (see annotated values). Regions enclosed by two adjacent iso-lines are colored uniformly (darker colors denote larger values). The input-kernel time constant is $\tau_{\text{in}} = 5$ ms. Similar results are obtained with different values of $\tau_{\text{in}} < \tau_{\text{out}}$. Parameter values: $\sigma = 6.25$ cm, $v = 0.25$ m/s, $L = 1$ m. $r_{\text{av}} = 0.4$ s^{-1}.

Finally, the inverse Fourier transform of Equation 5.92 reads

$$K^{\text{eq}}(t) = H \cdot \left[\left(\frac{1}{\tau_{\text{in}}} - \frac{1}{\tau_{\text{out}}} \right) \exp\left(-\frac{t}{\tau_{\text{in}}} \right) - \frac{\mu_{\text{out}}}{\tau_{\text{out}}} \exp\left(-\frac{t}{\tau_{\text{out}}/(1+\mu_{\text{out}})} \right) \right] \tag{5.94}$$

for $t \geq 0$ and $K^{\text{eq}}(t) = 0$ for $t < 0$. Equation 5.94 shows that the equivalent filter K^{eq} is a difference of two exponentials, similarly to the kernel K in Equation 5.2. Note however that the two exponentials are scaled differently as compared to the original filter K. Additionally, if the integral of the output kernel is negative, the integral of the equivalent filter is always positive (Equation 5.88 with $\omega = 0$).

To test whether spatially-periodic patterns could still emerge in this scenario, I compute the eigenvalue spectrum $\lambda(k)$ and the critical spatial frequency $k_{\max}$ by using Equations 5.31 and 5.32 with $K = K^{\text{eq}}$. Surprisingly, I find that typical grid scales (e.g., $k_{\max} > 2$ m^{-1}) are obtained for output-kernel time constants of the order of seconds, which seem biologically unrealistic (Figure 5.10). Therefore, I conclude that AHPs alone are not sufficient to generate grid-like patterns. Nevertheless, AHPs could still support structure formation by amplifying the effects of a band-pass filter that is already present at the input.

5.A.6. Numerical simulations

Model parameters and derived quantities are summarized in Tables 5.1 and 5.2.

Neural activity		
N		Number of synaptic inputs
r_0	[s^{-1}]	Baseline rate of the output neuron
τ_S	[s]	Adaptation kernel short time constant
τ_L	[s]	Adaptation kernel long time constant
μ		Adaptation kernel scaling parameter
Spatial exploration		
L	[m]	Side-length of the arena
v	[m/s]	Running speed of the virtual rat
σ_θ		Standard deviation of running directions
Input spatial tuning		
r_{av}	[s^{-1}]	Average input rate in the arena
σ	[m]	Width of the input receptive fields
M		Number of receptive fields per neuron (spatially-irregular inputs)
Synaptic plasticity		
η		Learning rate
τ_W	[s]	Decay time constant of the learning window W
W_{tot}	[s]	Integral of the learning window W
α		Multiplicative weight-normalization constant
β		Additive weight-normalization constant
Derived quantities		
a	[s^{-1}]	Multiplicative weight-normalization rate
b	[s^{-1}]	Additive weight-normalization rate
λ_{max}	[s^{-1}]	Maximal eigenvalue
w_{av}^{∞}		Average synaptic weight
τ_{av}	[s]	Weight normalization time scale
τ_{str}	[s]	Structure formation time scale

Table 5.1. | Model parameters

	Unit	Fig 5.4 Fig 5.5	Fig 5.6 (TL)	Fig 5.6 (TR)	Fig 5.6 (BL) Fig 5.7A1-4	Fig 5.6 (BR) Fig 5.7B1-4	Fig 5.8 Fig 5.9
N		900	3600	3600	3600	3600	3600
r_0	$[s^{-1}]$	10	4	4	4	4	4
τ_S	[s]	0.1	0.1	0.1	0.1	0.1	0.1
τ_L	[s]	0.16	0.16	0.35	0.16	0.35	0.16
μ		1.06	1.06	1.06	1.06	1.06	1.06
L	[m]	1	2	2	2	2	1
v	[m/s]	0.25	0.25	0.25	0.25	0.25	0.25
σ_θ		0.7	0.7	0.7	0.7	0.7	0.7
r_{av}	$[s^{-1}]$	0.4	0.21	0.085	0.3	0.1	0.8
σ	[m]	0.0625	0.045	0.045	0.0625	0.0625	0.0625
M		–	–	–	–	–	10
η		$2 \cdot 10^{-5}$	$5 \cdot 10^{-5}$	$5 \cdot 10^{-5}$	$5 \cdot 10^{-5}$	$5 \cdot 10^{-5}$	$5 \cdot 10^{-5}$
τ_W	[s]	0.05	–	–	–	–	–
W_{tot}	[s]	1	1	1	1	1	1
α		3.56	–	–	–	–	–
β		−8.78	–	–	–	–	–
a	$[s^{-1}]$	1.1	4	4	4	4	2.5
b	$[s^{-1}]$	0.49	0.69	0.28	1.23	0.31	2.8
λ_{max}	$[s^{-1}]$	1	0.90	0.82	0.80	0.85	1.75
w_{av}^{∞}		0.05	0.05	0.05	0.05	0.05	0.02
τ_{av}	[s]	$5.13 \cdot 10^3$	$1.44 \cdot 10^3$	$3.57 \cdot 10^3$	$8.12 \cdot 10^2$	$3.18 \cdot 10^3$	$1.42 \cdot 10^2$
τ_{str}	[s]	$5 \cdot 10^4$	$2.23 \cdot 10^4$	$2.42 \cdot 10^4$	$2.50 \cdot 10^4$	$2.36 \cdot 10^4$	$1.14 \cdot 10^4$

Table 5.2. | Default parameter values for the numerical simulations. See also Table 5.1 for short descriptions of the parameters. TL: top-left, TR: top-right, BL: bottom-left, BR: bottom-right. Note that in Figure 5.5A the learning rate η is varied from $2 \cdot 10^{-5}$ to $10 \cdot 10^{-5}$ and that in Figure 5.5B the virtual-rat running speed is sampled from an Ornstein-Uhlenbeck process with long-term mean $\bar{v} = v$.

Simulation of the detailed spiking model

The detailed spiking model (Figures 5.4 and 5.5) is simulated using the Brian2 simulation software (Stimberg et al., 2014). Neural and synaptic variables are integrated with a time step of 1 ms. The random walk of the virtual rat that is updated every 10 ms. The physical space explored by the virtual rat is discretized in 200^2 square bins.

Simulation of the averaged weight dynamics

The average weight dynamics (Equation 5.16) is integrated by using the forward Euler method with integration time step of 50 s (Figures 5.6-5.8). The input correlation matrix C is computed using Equation 5.54 for spatially-regular inputs, and using Equation 5.21 for spatially-irregular inputs.

Initialization of the synaptic weights

At the initial condition, the synaptic weights are normally distributed around the target normalization level $w_{\text{av}}^{\infty} = 5 \cdot 10^{-3}$. The standard deviation of the distribution is 10^{-4} for the spiking simulations and 10^{-3} for the average weight dynamics.

5.A.7. Data analysis

Grid properties

I compute the grid spatial scale from the two-dimensional Fourier amplitude of the grid pattern. I estimate the radial amplitude profile by averaging over the angular dimension. I then define the grid scale as the frequency where the amplitude profile has a global maximum.

The grid orientation is estimated from the spatial auto-correlogram of the grid pattern. I detect the peak closest to the center in the first quadrant of the auto-correlogram. I then define the grid orientation as the angle between the detected peak and the horizontal axis.

I define the grid spatial phase as the position of the closest peak to the center in the cross-correlation between the grid pattern and a reference grid at the same scale.

Gridness score

I estimate the gridness score similarly to Langston et al. (2010). First, I compute the spatial auto-correlogram of the weight (or firing-rate) pattern and I retain only points within a ring of outer radius R_i and inner radius $R_i/2$. I then compute the gridness

score g_i as

$$g_i := \frac{1}{2}\left[\rho_i(60) + \rho_i(120)\right] - \frac{1}{3}\left[\rho_i(30) + \rho_i(90) + \rho_i(150)\right] \tag{5.95}$$

where $\rho_i(\varphi)$ is the Pearson's correlation coefficient between the original ring (of outer radius R_i) and the same ring rotated by φ degrees. The final gridness score is defined as the maximum g_i by varying the outer radius R_i between $0.7/k_{\max}$ and $2.5/k_{\max}$ where $k_{\max}$ is the spatial frequency of the pattern.

Estimation of output firing-rate maps

The output firing-rate maps Ψ^{out} in Figure 5.8 B are computed as follows:

$$\Psi^{\text{out}}(\mathbf{x}) = r_0 + \int \mathrm{d}\mathbf{y}\, K_{\text{sp}}(|\mathbf{y}|) \sum_{i=1}^{N} w_i \Psi_i^{\text{in}}(\mathbf{x} - \mathbf{y}) \tag{5.96}$$

where r_0 is the baseline firing rate, w_i are the synaptic weights at the end of the simulation, Ψ_i^{in} are the input spatial maps, and K_{sp} is the equivalent adaptation kernel in space (Equation 5.31). The convolution with the filter K_{sp} accounts for the average effect of the temporal kernel K on the output firing rate.

Chapter 6

Inheritance and amplification of grid-cell activity

Here, I study how network dynamics affect grid-cell patterns. In particular, I seek to determine how excitatory feed-forward and recurrent circuits may contribute to the inheritance and amplification of grid-cell tuning. A subset of the results in this chapter (Section 6.2, feed-forward amplification) have been included in a peer-reviewed journal article which I co-authored (Winterer et al., 2017).

6.1. Introduction

Grid cells are neurons of the medial entorhinal cortex (MEC) that are tuned to the animal's position in the environment and whose firing fields tile the space with a regular triangular pattern. Since their discovery, grid cells have attracted considerable research attention, due to the richness of their responses, and because they are thought to underlie high-level cognitive functions, such as spatial navigation and spatial memory (e.g., McNaughton et al., 2006; Mathis et al., 2012a; Ólafsdóttir et al., 2016; Gil et al., 2018; Tennant et al., 2018, Chapter 8). Nonetheless, it remains to date unclear how grid-cell activity is formed and how it is affected by synaptic connectivity within the cortex (Rowland et al., 2016)

Grid cells have been found in all principal-cell layers of the MEC (Sargolini et al., 2006), and they are most abundant in the superficial layers (II and III). Grid-cell activity in layer II has been reported in both stellate and pyramidal neurons (Schmidt-Hieber and Häusser, 2013; Domnisoru et al., 2013; Tang et al., 2014; Sun et al., 2015)—the two main excitatory-cell types in the area (Klink and Alonso, 1997). Interestingly, cell type and cortical layering also define local microcircuits in the MEC with stereotypical feed-forward pathways, e.g., layer II stellate cells project to pyramidal cells in layer V (Sürmeli et al., 2015) and receive direct synaptic input from pyramidal cells in layer II and III (Fuchs et al., 2016; Winterer et al., 2017, see also Section 2.2.1 and Figure 2.2).

Within local populations, principal neurons are also recurrently connected, i.e., collaterals exist among layer III and layer V principal neurons (Dhillon and Jones, 2000; Winterer et al., 2017), and among layer II stellate and pyramidal cells (Fuchs et al., 2016; Winterer et al., 2017; Schmidt et al., 2017).

Although *in vitro* studies cannot assess the spatial tuning of the connected neurons, three lines of evidence indicate that grid cells could be recurrently coupled. First, grid-cell activity is organized in distinct functional modules characterized by a shared grid scale and orientation (Hafting et al., 2005; Stensola et al., 2012). Second, cells of the same module react in concert to external manipulations of the environment (Stensola et al., 2012; Yoon et al., 2013). Third, the spiking activity of nearby grid cells is temporally correlated beyond what is expected from simple firing-rate covariation (Dunn et al., 2015; Tocker et al., 2015). It was proposed that grid-cell coupling could arise through disynaptic inhibitory circuits (Couey et al., 2013). However, this requires a structured inhibitory connectivity that depends on the spatial phases of the excitatory grids—a scenario that is at odds of available empirical data (Buetfering et al., 2014).

In summary, experimental evidence suggests that grid cells are found in distinct neuronal populations of the MEC, and that these populations are embedded in excitatory feed-forward and recurrent circuits. But what could be the effects of such connections on grid-cell activity? Would they disrupt or rather amplify grid patterns? Do grids originate in different populations independently, or can they be inherited through feed-forward circuits? And how are grids affected by recurrent network dynamics?

In this chapter, I address these questions from a theoretical standpoint. In Section 6.2, I focus on feed-forward networks, and I show that, under certain conditions, grids can be inherited and even amplified by feed-forward projections. Additionally, I demonstrate that a connectivity supporting these functions can self-organize in a Hebbian framework. I then study how grid patterns are affected by recurrent dynamics (Section 6.3). I show that grid amplification can also be obtained in recurrent circuits, and I relate my findings to continuous-attractor models of grid-cell firing. Finally, I propose a minimal mathematical model for the amplification of grid-cell activity on linear tracks (Section 6.4). In this model, I analytically quantify the amount of amplification that a network can achieve as a function of the properties of the input tuning and the connectivity between the neurons.

6.2. Dynamics of grid-cell activity in feed-forward networks

First, I study the dynamics of grid-cell activity in feed-forward networks. For clarity, I study inheritance and amplification separately: in Section 6.2.1, I focus on the inheritance problem; in Section 6.2.2, I extend the results to an amplification framework.

6.2.1. Feed-forward inheritance of grid-cell activity

I start by assessing whether grid patterns can be inherited by feed-forward projections. To this end, I model the activity of a population of input grid cells in an upstream region that project to a population of output cells in a downstream region (Figure 6.1A1, detailed methods in Section 6.6). Input grids have similar scales and orientations, but distributed spatial phases (Figure 6.1A2)—similar to experimentally observed grids in a single module (Hafting et al., 2005; Stensola et al., 2012). Aside from synaptic excitation, output cells receive inhibition from local interneurons, and inhibition is spatially untuned. This is consistent with the facts that inhibitory connectivity is frequent within the MEC (Couey et al., 2013; Fuchs et al., 2016) and that inhibitory spatial tuning is typically broad (Buetfering et al., 2014).

I first consider the case in which the excitatory feed-forward connectivity is sparse and unstructured, alike to random synaptic wiring. In this scenario, each output cell is connected to input grids with random spatial phases (Figure 6.1B1), and, as a result, spatially-irregular patterns emerge at the output (Figure 6.1B2). Even though random connectivity could occasionally generate periodic outputs (Figure 6.1B2-3), the output rates distribute in a narrow dynamic range, which make such responses virtually indistinguishable from background noise. Therefore, random excitatory connectivity is unlikely to support the inheritance of grid-cell activity in the MEC.

By contrast, grid-field inheritance could result from feed-forward projections that are properly structured. That is, output grids could be obtained by integrating the activity of input grids with similar spatial phases. Additionally, such a connectivity structure could spontaneously emerge in a Hebbian framework.

To test these hypotheses, I simulate the development of the excitatory synaptic weights as a virtual rat explores a quadrangular enclosure. The synaptic plasticity rule is Hebbian, i.e., the weight changes are linearly related to the product of the input and output activities. To prevent an unbounded synaptic growth, the individual connection strengths are constrained to a fixed range (0 and $w_{\max}$), and the total synaptic weight is kept constant throughout the simulation; see Section 6.6.2 for detailed methods and (Turrigiano and Nelson, 2004) for experimental evidence on synaptic-weight homeostasis.

Figure 6.1C shows the simulation results for one example output neuron in the network (see also Cell 1 in Figure 6.1B, D). Initially, the cell is driven by a random set of inputs (Figure 6.1C1-2 at t = 0 min), and the corresponding output firing-rate map is spatially irregular (Figure 6.1C3 at t = 0 min). With experience, however, a structure emerges in the synaptic weights such that inputs with nearby grid phases are potentiated and all others inputs are depressed (Figure 6.1C1-2 at t = 50, 100, 200 min). As a result, a nearly-regular grid pattern emerges at the output (Figure 6.1C3 at t = 50,

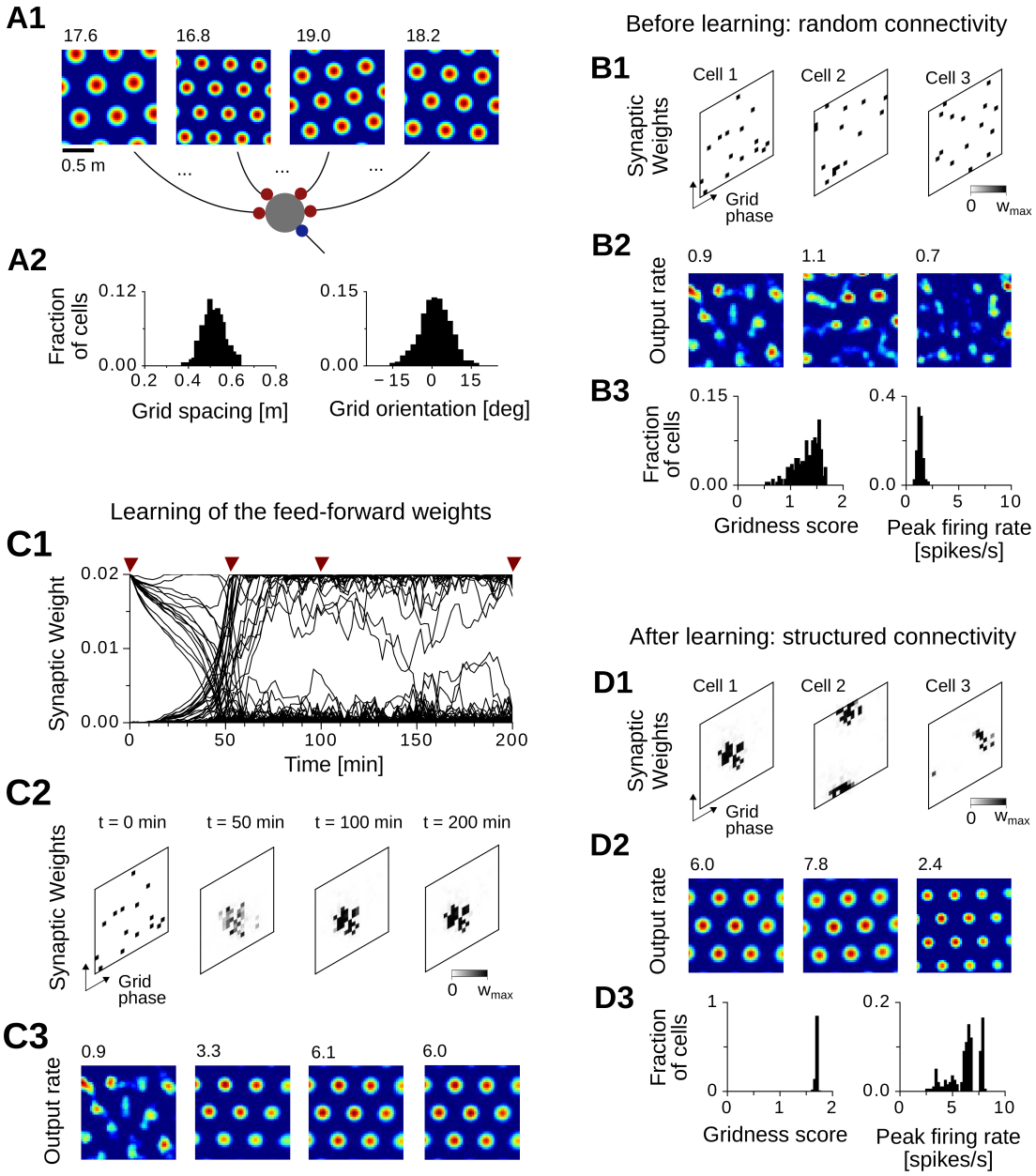

Figure 6.1. | Grid-field inheritance in feed-forward networks. A) Model schematic and input tuning. A1) One cell (gray disk) in a target neuronal population receives excitation (red disks) from an upstream population of grid cells (example spatial patterns at the top row, dark blue: 0 spikes/s, dark red: peak rate in spikes/s at the top-left corner) and inhibition (blue disk) from spatially-untuned interneurons. Input grids have similar spacing and orientation (distributions in A2), but different spatial phases. The excitatory synaptic weights are initially random (binary), and they are plastic according to a Hebbian learning rule. A2) Distributions of the spacings (left) and orientations (right) of the input grids ($N_{\text{in}} = 400$). **B)** Synaptic weights and output firing-rate maps before learning. B1) Excitatory weights of three example output cells in the target population. Nearby pixels in the weight maps correspond to inputs with similar grid phases. B2) Output firing-rate maps resulting from the weights at the top row. B3) Distribution of the gridness scores (left) and peak rates of the output firing-rate maps ($N_{\text{out}} = 200$). **C)** Development of the synaptic weights and output firing-rate maps during learning. C1) Weights evolution for one example output cell in the target population (see Cell 1 in panels B and D). C2) Snapshots of the synaptic weights at four time steps during learning (see B1 and red triangles in C1). C3) Output firing-rate maps resulting from the weights in C2. **D)** Synaptic weights and output firing rate maps after learning (see panel B). See Section 6.6 for further details and parameter values.

100, 200 min).

The spatial phase of the output grid depends on the initial state of the synaptic weights, which was random. Therefore, with an initially random synaptic wiring, the output neurons develop grids with distributed spatial phases (Figure 6.1D1-2). These results show that grid-cell activity can be inherited by structured feed-forward projections (Figure 6.1D3), and that such projections can self-organize via a simple Hebbian rule.

6.2.2. Feed-forward amplification of grid-cell activity

Next, I test whether feed-forward networks could even improve (or amplify) the spatial regularity of noisy grids. From a theoretical point of view, grid-pattern amplification would be particularly desirable if grids originated via a noisy single-cell process, such as the one proposed in Chapter 5. Additionally, grid patterns could be distorted by spatially-irregular inputs, which are often found in the MEC (Diehl et al., 2017). In this scenario, studying amplification in feed-forward circuits could help to identify functionally upstream and downstream grid-cell populations, possibly pointing to the locus of origin of grid-cell activity.

I study grid-pattern amplification in the same feed-forward network described in Section 6.2.1. Here, however, the feed-forward inputs are only weakly grid-tuned. I consider two input scenarios: grids with jittered firing-field centers (Figure 6.2A1) and grids distorted by additive noise (Figure 6.2B1). In the first case, inputs are generated by randomly shifting the firing-field locations of regular triangular grids. In the latter case, regular grids are superimposed to a blanket of spatial noise, which is smooth in space but uncorrelated across neurons (see Section 6.6.5 for detailed methods). In both cases, the underlying regular grids are co-modular and the degraded patterns have low gridness scores (Figure 6.2A2, B2).

Alike to Section 6.2.1, the excitatory synaptic weights are initially random and they are subject to a Hebbian learning rule as a virtual rat explores the environment (see Sections 6.6.2 and 6.6.6 for detailed methods). In this case, despite the degraded inputs, a clear structure develops in the synaptic weights and more regular grids emerge at the output (Figure 6.2A3, B3 and Figure 6.2A4, B4). Indeed, Hebbian synaptic plasticity is sensible to even small correlations at the input ensemble, albeit large distortions of the individual input patterns. Such correlations are then amplified by the network dynamics, increasing the spatial periodicity of the output maps.

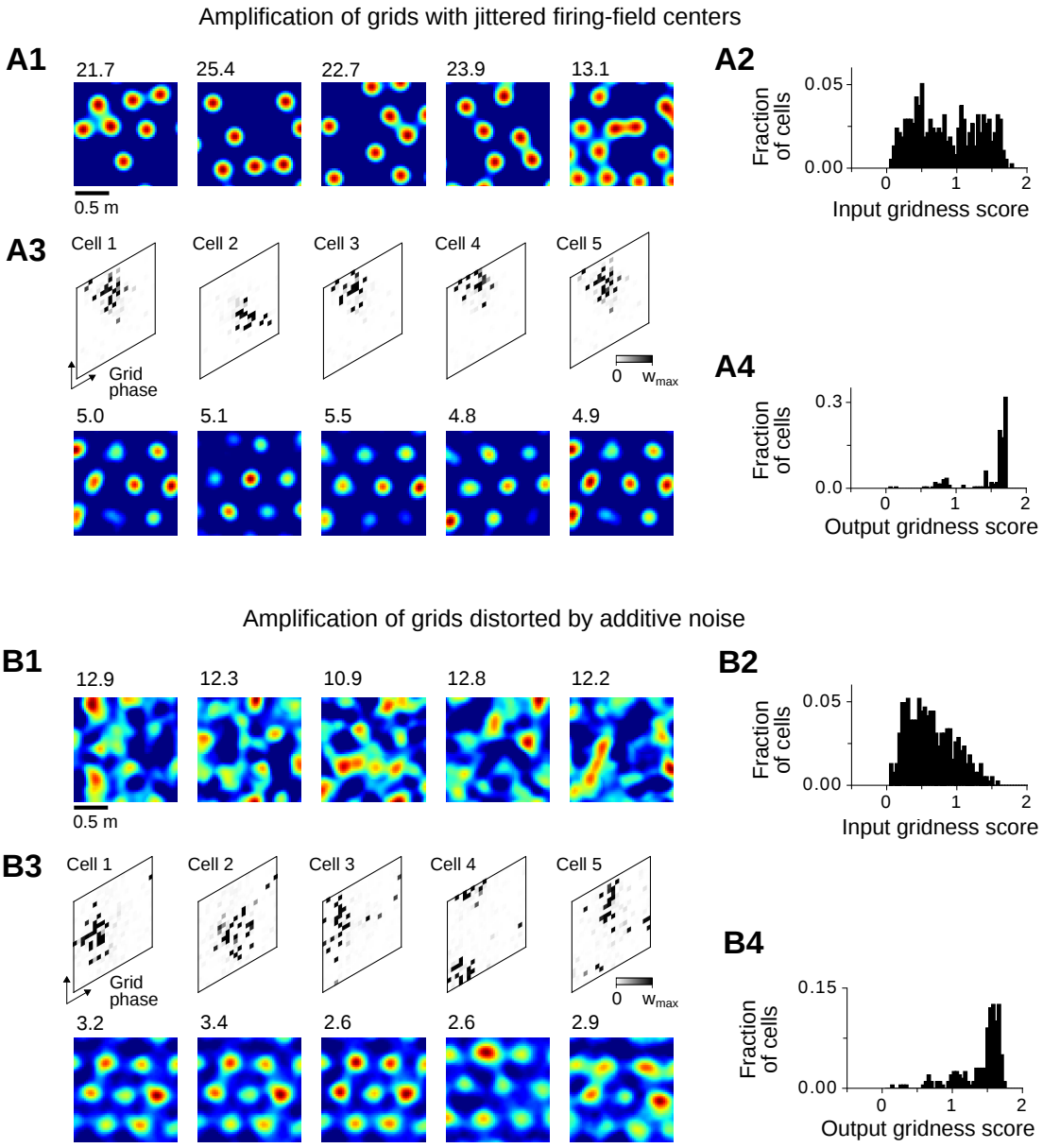

Figure 6.2. | Grid amplification in feed-forward networks. A) Amplification of grids with jittered firing-field centers. A1) Example input firing-rate maps (dark blue: 0 spikes/s, dark red: peak rate in spikes/s at the top-left corner). Input maps were generated by jittering the firing-field locations of regular triangular grids. The underlying grids had similar grid spacing and orientation but different phases (see examples and distributions in Figure 6.1A2). A2) Distribution of the input gridness scores ($N_{\text{in}} = 400$). A3) Synaptic weights (top row) and firing-rate maps (bottom row) of five example output cells. The synaptic weights were learned directly from the inputs starting from a random initial condition (see also Figure 6.1C). Weights are arranged according to the spatial phase of the underlying regular grids. A4) Distribution of the output gridness scores ($N_{\text{out}} = 200$). **B)** Amplification of grids distorted by additive noise. Input patterns were generated by adding random spatial maps to regular grids with similar spacing and orientation but different phases (see panel A). See Section 6.6 for further details and parameter values.

6.3. Dynamics of grid-cell activity in recurrent networks

I now study the dynamics of grid-cell activity in recurrent networks. First I show that recurrent circuits—similarly to feed-forward networks—can also improve the regularity of grid patterns.

I model the activity of a recurrent network including both excitatory and inhibitory cells (Figure 6.3A). The two neuronal populations are recurrently connected locally and they are mutually coupled via feed-forward projections. The external drive to the network is provided by spatially-tuned feed-forward inputs that project exclusively to the excitatory population.

I assume that the total feed-forward input to each excitatory cell is weakly grid-tuned as a function of space, i.e., it is a regular triangular grid in space distorted by additive noise (Section 6.6.5, see also Figure 6.2B1). Alike to a grid-cell module, the underlying input grids have similar spacing and orientations, and uniformly-distributed spatial phases. Therefore, the total feed-forward input to each excitatory cell is a noisy grid with a particular spatial phase, which I call the *preferred phase* of the neuron.

To mimic a biologically-realistic network, the excitatory neurons outnumber the inhibitory neurons by a factor of four, i.e., I model the activity of $N_{\mathrm{E}} = 900$ excitatory and $N_{\mathrm{I}} = 225$ inhibitory neurons. Additionally, in line with experimental data (Fuchs et al., 2016; Winterer et al., 2017), the excitatory connectivity is sparser than the inhibitory one: each neuron in the network receives input from $\sim 30\%$ of the inhibitory cells, but only $\sim 6\%$ of the excitatory ones. The synaptic weights of both excitatory and inhibitory connections are initially random.

6.3.1. Self-organization of the excitatory connections

To amplify grid spatial patterns, recurrent (like feed-forward) circuits require structured synaptic connections. Here, I hypothesize that such a structured recurrent connectivity could emerge in the excitatory weights via synaptic plasticity.

To test this hypothesis, I simulate the development of the excitatory weights under a simple Hebbian rule. The rule strengthens inputs that correlate with the output firing, but keeps the overall connectivity rates constant and the individual synaptic weights bounded (see Section 6.6.4 for detailed methods). The weight dynamics is simulated while a virtual rat explores the environment for 500 s.

Figure 6.3B shows the results of such a learning process. Before learning (Figure 6.3B1), the excitatory weights are random, and, accordingly, there is no relation between the *preferred phase* of an excitatory cell (pink dot) and the phases of the cells it receives input from (black pixels). After learning, however, a structure emerges in the synaptic weights such that only cells with similar preferred phases are excitatorily

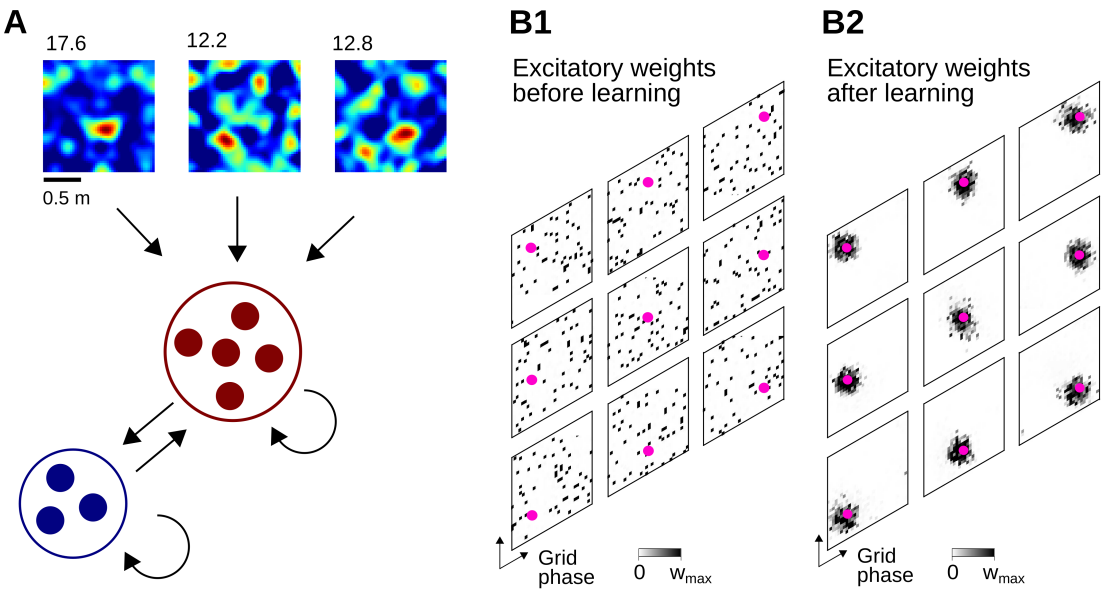

Figure 6.3. | Grid amplification in recurrent networks: model schematic and synaptic connectivity. A) Model schematic. A recurrent network of $N_E = 900$ excitatory neurons (red disks) receive excitation from spatially-tuned cells (three example firing-rate maps in the top row, dark blue: 0 spikes/s, dark red: peak rate in spikes/s at the top-left corner) and inhibition from $N_I = 225$ interneurons (blue disks). The black arrows indicate feed-forward and recurrent connections between neuronal populations. The total feed-forward input to each excitatory cell is a noisy grid (i.e., a regular grid distorted by additive noise). **B)** Recurrent excitatory connectivity. The recurrent excitatory weights are initially random and they are plastic according to a Hebbian learning rule. The plots show the excitatory connections to nine example principal cells in the network before learning (B1) and after learning (B2). Each synaptic weight map (small rhombus) shows the input connections to a cell whose preferred phase is denoted by the pink dot. Nearby pixels in a synaptic weight map correspond to cells with similar phases (see also Figures 6.1 and 6.1). See Section 6.6 for further details and parameter values.

connected (Figure 6.3B1).

6.3.2. Recurrent amplification of grid-cell activity

Next, I simulate the output activity of the network with the connectivity that developed after learning. Note that the inhibitory connections are still random, i.e., an inhibitory neuron receives input from (and provides output to) a random set of cells. I also assume that the recurrent dynamics is much faster than the movements of the virtual rat, which is reasonable for typical values of synaptic time constants ($2 - 5$ ms), membrane time constants (~ 10 ms), and average exploration speeds of laboratory rats ($10 - 30$ cm/s). With this assumption, I simulate the network activity at all spatial locations independently, i.e., without explicitly modeling a virtual-rat random walk in the environment.

Figure 6.4A1 shows the evolution of the output activity for one example excitatory cell in the network. Initially, the activity is dominated by the feed-forward inputs, and the output firing is spatially irregular ($t = 2 - 10$ ms). Within tens of milliseconds, however, the recurrent input is integrated by the network, and the output firing becomes more spatially regular (A1, $t > 20$ ms). With a network-integration time constant of $\tau = 10$ ms, the output firing approaches a steady state after about 60 ms (Figure 6.4A2). This suggests that recurrent amplification is fast enough to occur within a single cycle of the theta rhythm ($\sim$100-150 ms)—an oscillation that strongly paces network activity in the MEC.

The recurrent dynamics increases the spatial periodicity of all excitatory cells in the network (Figure 6.4B1 and B2). By contrast, only a very weak periodic tuning is observed in the inhibitory neurons (Figure 6.4C1 and C2). This result is consistent with experimental data showing spatially-broad inhibitory tuning in the MEC (Buetfering et al., 2014). Note, however, that the gridness scores of the simulated inhibitory maps are considerably higher than experimentally reported (Buetfering et al., 2014). Indeed, in spite of a random inhibitory connectivity, interneurons often inherit weak spatial periodicity from the excitatory population, which might be due to the small number of neurons in the simulations shown (see also Figure 6.1B for a similar effect with excitatory neurons). Yet, in line with the experiments, interneurons typically fire throughout the environment, and the dynamic range of their responses is very narrow (Figure 6.4C1).

6.3.3. Relation between amplification and attractor models

Recurrent amplification of grid-cell activity crucially relies on structured feedback connections, i.e., a recurrent connectivity where grids with nearby phases are synaptically coupled. Notably, such a connectivity is also assumed by continuous attractor network (CAN) models for the origin of grid-cell patterns (e.g., McNaughton et al., 2006; Fuhs and Touretzky, 2006; Guanella et al., 2007; Burak and Fiete, 2009, see also Section 4.1). Here, I point out differences and similarities between CAN and amplification models, and I show why recurrent connectivity alone is not sufficient to generate periodic tuning.

To this end, I visualize the activity of the excitatory neurons at two different levels: the single-cell level and the population level. At the single-cell level, the activity of a particular neuron is shown for all positions of the virtual rat in the environment (Figure 6.5A1, A4, B1, B4). At the population level, the activity of all neurons is shown for a particular virtual-rat position in the environment (Figure 6.5A2, A3, B2, B3). Note that, at the population level, a choice has to be made regarding the arrangement of the neuronal activity on a single map. Because the recurrent excitatory connectivity

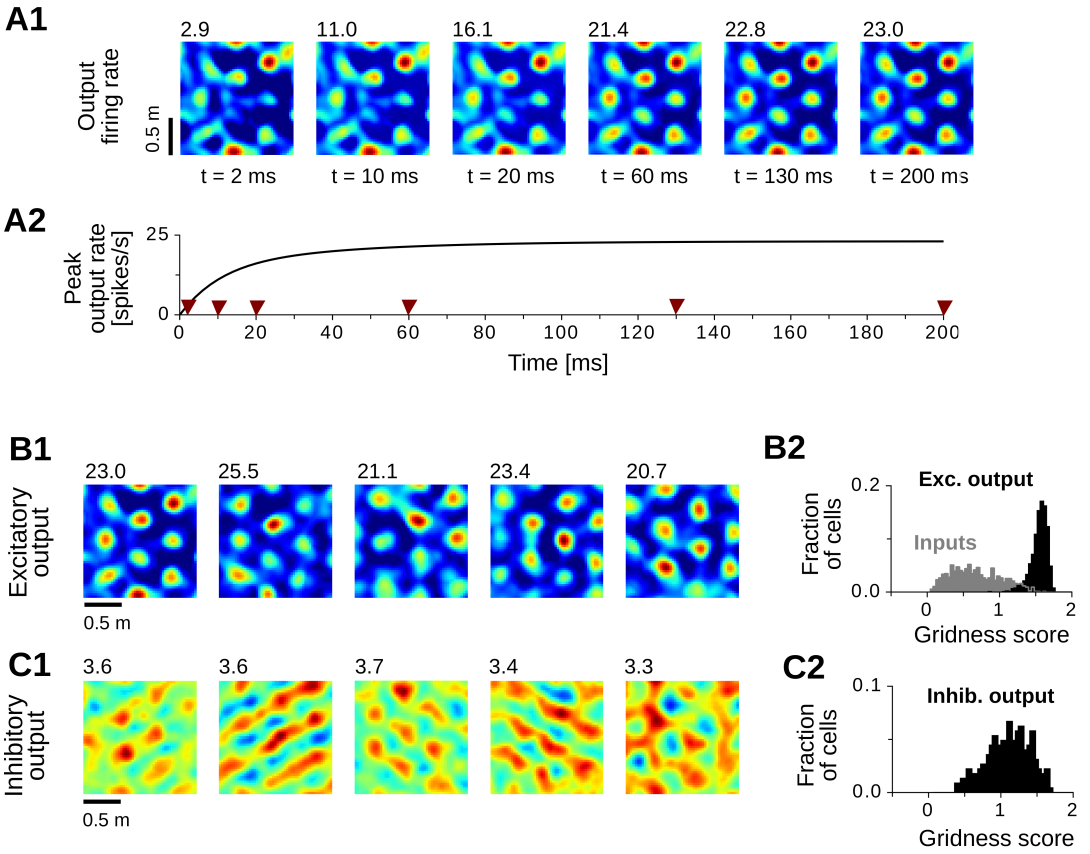

Figure 6.4. | Grid amplification in recurrent networks: temporal dynamics and steady-state patterns. **A)** Temporal dynamics of grid-pattern amplification. A1) Evolution of the output firing-rate map of one example excitatory cell. The network activity is integrated for each spatial location independently for 200 ms. The network integration time constant is $\tau = 10$ ms. The output firing-rate map is shown at six time points, i.e., t = 2, 10, 20, 60, 130 and 200 ms (see also red triangles in A2). Dark blue: 0 spikes/s, dark red: peak rate in spikes/s at the top-left corner. A2) Evolution of the peak firing rate of the spatial maps at the top row. Note that after about 60 ms the output activity approaches as steady state. **B)** Steady-state spatial tuning of the excitatory neurons. B1) Output firing-rate maps of five example excitatory neurons after 200 ms of integration. B2) Gridness score distribution of the steady-state spatial maps of the excitatory neurons ($N_E = 900$, black bars) and of the total feed-forward inputs ($N_E = 900$, gray bars). **C)** Steady-state spatial tuning of the inhibitory neurons ($N_I = 225$). See panel B for details. Although inhibitory spatial maps show weak spatial periodicity, the dynamic range of their responses is very narrow. See Section 6.6 for further details and parameter values.

is structured, it is useful to arrange the neurons in a way that mirrors the topology of the underlying network. Therefore, in the population maps of Figure 6.5, neurons are arranged according to the phases of the input grids that structured the excitatory weights during learning. Accordingly, nearby pixels in the population maps correspond to strongly excitatorily connected cells in the network. Note that this arrangement bears no relation to the physical location of the neurons in the anatomical tissue.

A key difference between CAN and amplification models resides in the spatial tuning of the feed-forward inputs, which are assumed to be weakly grid-tuned in the amplification model (Figure 6.6A1), but spatially aperiodic in attractor models (Figure 6.6B1). I study these two scenarios separately.

First, I consider a network that receives weakly grid-tuned feed-forward inputs, as assumed in the amplification scenario. Because the feed-forward inputs are weakly grid-tuned, and these inputs were used to learn the excitatory weights, co-active cells in the network are also likely to be excitatorily coupled. In other words, the input spatial tuning matches the topology of the underlying network. As a consequence, an activity bump emerges at the population input (A2). The bump is formed by co-active cells in a particular location of the environment (e.g., Pos 1 in A2). As the virtual rat moves to a different location, a new set of cells is active, and the population bump moves across the network (see Pos 1-5 in A2). Additionally, owing to the spatial periodicity of the inputs, the bump reoccurs at the same location in phase space (rhombus) after the rat has moved the distance of one grid period (along one of the grid axes) in physical space (square, compare Pos 1 and Pos 5 in A2). In other words, the position of the population bump in phase space tracks the position of the virtual rat in physical space. Importantly, the recurrent connections sharpen this population bump without affecting its location (A3) and, as a result, more regular grids emerge at the single-cell output (compare A1 and A4).

In CAN models, instead, the feed-forward inputs provide noisy excitation to all neurons, but their activity is spatially aperiodic (e.g. Fuhs and Touretzky, 2006; Guanella et al., 2007; Burak and Fiete, 2009). To mimic this scenario, I now assume that the total feed-forward input to each neuron is a random spatial map (one example in B1). In this case, the input tuning bears no relation to the topology of the recurrent connections and, accordingly, there is no activity bump at the population input (B2).

Yet structured recurrent connectivity—if strong enough—can generate an activity bump at the population output by amplifying the noise at the feed-forward input (B3). The location of this bump, however, depends on the input tuning, which is now random and spatially aperiodic. As a consequence, the bump randomly shifts in phase space even for straight virtual-rat trajectories in physical space (one example in B3, Pos 1-5). In other words, a direct link between the location of the bump and the movements of the virtual rat is completely lost. Accordingly, grid-like patterns do not emerge at

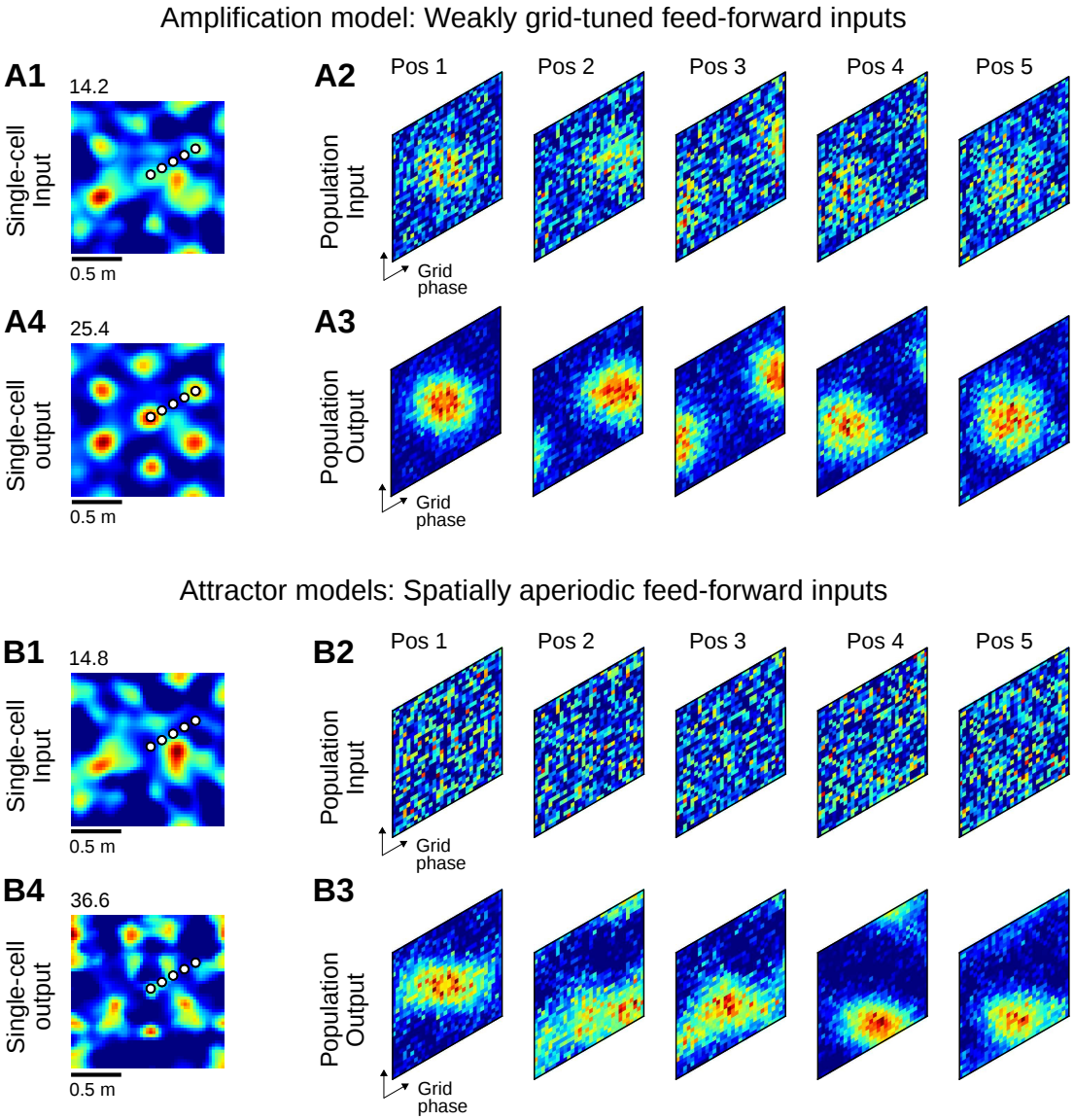

Figure 6.5. | Relation between amplification and attractor models. **A)** Network dynamics in the amplification model. The total feed-forward input to each excitatory neuron is a noisy grid (one example in A1, see also Figures 6.2B1 and 6.3A). The excitatory recurrent connectivity emerged via Hebbian learning (Figure 6.3B2). The square plots show the total feed-forward input (A1) and the steady-state output (A4) of a particular cell in the network (single-cell level). Dark blue: 0 spikes/s, dark red: peak firing rate in spikes/s at the top-left corner. The rhomboidal plots show the total feed-forward input (A2) and the steady-state output (A3) of all excitatory neurons for five virtual-rat position in the environment (Pos 1-5, see white dots in A1 and A4). Dark blue: 0 spikes/s, dark red: peak firing rate in the map. In the population plots, neurons are arranged according to the input grid phases that shaped the recurrent connections during learning. The network activity shown is the result of the network dynamics simulated for any location in the environment for 200 ms. **B)** Network dynamics in continuous attractor models. The total feed-forward input to each excitatory neuron is a random and aperiodic spatial map (one example in B1). The excitatory recurrent connectivity has the same structure as in A (Figure 6.3B2), but the synaptic weights are 50% larger. See Section 6.6 for further details and parameter values.

the single-cell output (B4). This case illustrates that recurrent connectivity alone is not sufficient to generate periodic grids at the single-cell level.

Yet CAN models can generate single-cell grids even in the absence of spatially-periodic inputs. For this to happen, CANs assume an additional network mechanism that shifts the population bump in phase space according to the animal's movements in physical space. Such a mechanism requires an extra set of inputs that inform the cells about the animal's current speed and direction of movement (see Section 4.1 for details). In this case, the location of the population bump in phase space becomes completely uncoupled from the feed-forward inputs, i.e., it depends only on the initial condition and on the trajectory of the animal. Therefore, by integrating instantaneous velocity signals, the population bump tracks the animal's location in the environment.

If this integration was carried out without errors, spatially-periodic maps would emerge at the single-cell output due to the periodic recurrent connections. CANs, however, are inherently noisy velocity integrators, and—without anchoring to the physical space—integration errors accumulate rapidly. Therefore, I argue that recurrent dynamics and velocity inputs alone are unlikely to generate grid-cell activity *de novo*. Instead, recurrent connections could serve to amplify weakly-periodic inputs and—with the help of self-motion signals—even sustain grid-like firing when the feed-forward tuning is temporally lost.

6.4. Mathematical results on grid-pattern amplification

To elucidate the principles underlying the recurrent amplification of grid-cell activity, and to quantify this amplification for different input scenarios, I now study a more abstract mathematical model of the same phenomenon. This model captures the salient features of the amplification dynamics, is analytically tractable, and can be understood intuitively.

6.4.1. Model of neural activity

I model the activity of a recurrent network with weakly grid-tuned feed-forward inputs as a virtual rat runs on a linear track of length L. All input grids have the same spatial frequency, but different spatial phases ϕ, similarly to what is empirically observed within a grid-cell module (Stensola et al., 2012). Therefore, all neurons in the network are labeled according to the spatial phase ϕ of their total feed-forward input h. I call this phase the *preferred spatial phase* of a neuron (Figure 6.6A, B).

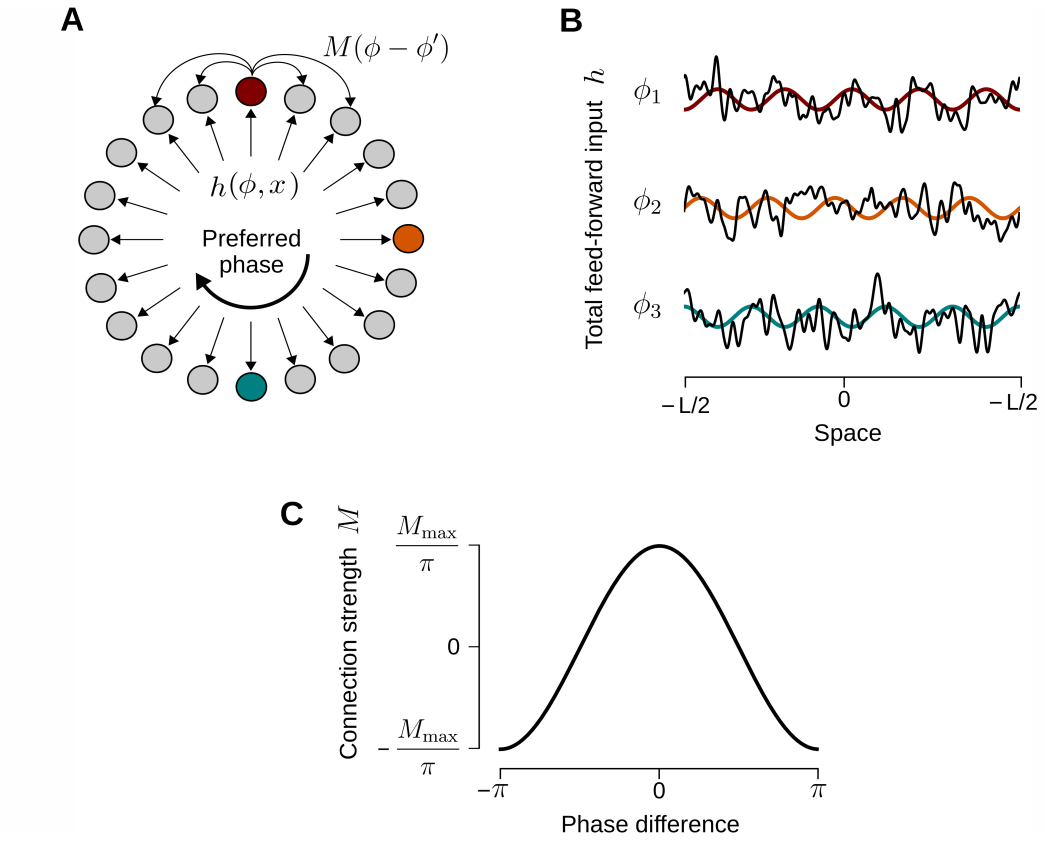

Figure 6.6. | Model of grid-pattern amplification on linear tracks. A) Cartoon of the modeled neuronal network. The total feed-forward input h to each neuron is weakly grid-tuned as a function of space (see examples in panel B). Neurons (disks) are arranged according to the grid phase ϕ of the feed-forward inputs. Colored disks indicate neurons with phases $\phi_1 = 0$, $\phi_2 = \pi/2$, and $\phi_3 = \pi$ (see panel B). The recurrent connectivity is defined by the function M, which depends only on the phase difference $\phi - \phi'$ between two neurons (see panel C). **B)** Feed-forward inputs to the network. The feed-forward inputs are weakly grid-tuned as a function of space on a linear track of length L. The black traces show the total feed-forward input h to three example neurons with preferred phases $\phi_1 = 0$, $\phi_2 = \pi/2$, and $\phi_3 = \pi$. The colored traces indicate the underlying grid-tuning function g with phases ϕ_1, ϕ_2, and ϕ_3 (Equation 6.3, see colored disks in panel A). **C)** Recurrent-connectivity function. The parameter $M_{\max} > 0$ scales the maximal connection strength in the network (Equation 6.2).

Output activity

I assume that each neuron integrates its input linearly, and that all preferred phases ϕ are evenly sampled in the range $[-\pi, \pi]$. Therefore, in the limit of a large number of neurons, the output activity of the network obeys

$$\tau \frac{\mathrm{d}}{\mathrm{d}t} v(\phi, x) := -v(\phi, x) + h(\phi, x) + \int_{-\pi}^{\pi} \mathrm{d}\phi' \, M(\phi - \phi') v(\phi', x) \tag{6.1}$$

where x is the position of the virtual rat on the track, $h(\phi, x)$ is the total feed-forward input to neuron ϕ, the function $M(\phi - \phi')$ defines the connection strength from neuron ϕ' to neuron ϕ, and $\tau > 0$ sets the integration time constant of the network. Note that the connectivity function M is translation invariant, i.e., it depends only on the phase difference $\phi - \phi'$.

Recurrent connectivity

I assume that neurons with similar phases are strongly excitatorily connected, whereas neurons with dissimilar phases are weakly excitatorily connected or inhibited:

$$M(\phi - \phi') := \frac{M_{\max}}{\pi} \cos(\phi - \phi') \tag{6.2}$$

where $M_{\max} > 0$ sets the maximal connection strength in the network, and the factor $1/\pi$ is for mathematical convenience (Figure 6.6C). Such a connectivity resembles the one emerging via Hebbian synaptic plasticity in the more detailed network models of Sections 6.2 and 6.3 (Figures 6.1 and 6.3). Note, however, that the connectivity function M assumes both positive and negative values, i.e., it models the compound effects of both excitatory and inhibitory synapses.

Input spatial tuning

The total feed-forward input h to a neuron with preferred phase ϕ is

$$h(\phi, x) := g(\phi, x) + \xi(\phi, x) \quad \text{with} \quad g(\phi, x) := B \cos(2\pi f x + \phi) \tag{6.3}$$

where g is the input grid signal, ξ is the input noise, and the parameter $B > 0$ scales the signal amplitude.

The input noise $\xi(\phi, x)$ is a two-dimensional Gaussian random field with mean $\mu_\xi = 0$ and autocorrelation $C_\xi(\tau_\phi, \tau_x)$. I consider the general case in which correlations may exist both in space and across neurons. Noise correlations in space could stem from slowly-varying sensory inputs. Noise correlations across neurons could stem from convergent feed-forward connectivity upstream to the modeled network. I further

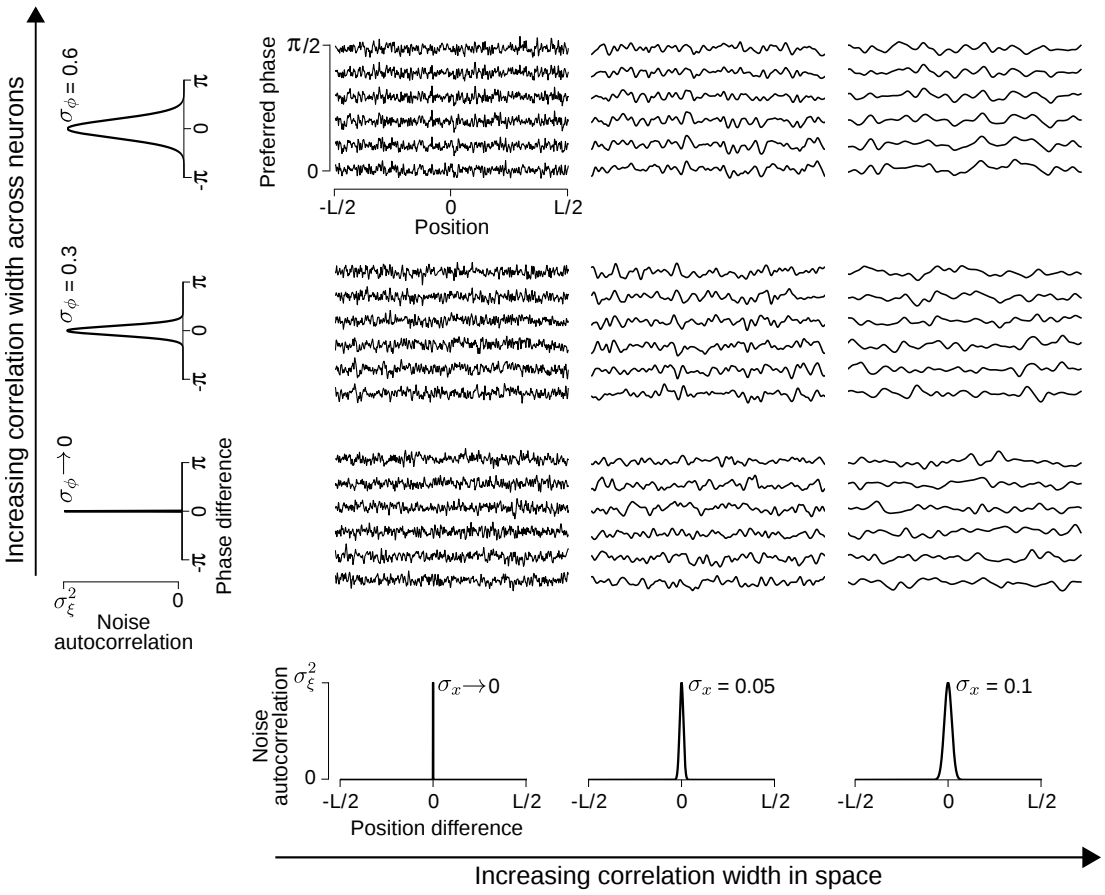

Figure 6.7. | Model of the input noise. Example traces of the input noise ξ (Equations 6.3 and 6.5) for three correlation lengths in space (σ_x, columns) and three correlation lengths across neurons (σ_ϕ, rows). The bell-shaped curves in the outer frame show the noise autocorrelation function C_ξ in space (bottom) and across neurons (left, Equation 6.5). For each parameter configuration, six example traces of the noise are shown (inner panels). Each trace depicts the noise to a neuron with a particular preferred phase. For illustration purposes, preferred phases are evenly sampled in the range $[0, \pi/2]$. By increasing the correlation length σ_x (from left to right), the noise becomes increasingly smooth in space. By increasing the correlation length σ_ϕ (from bottom to top), more and more neurons receive correlated input. The plots with $\sigma_\phi \to 0$ were generated by drawing the noise independently across neurons. The plots with $\sigma_x \to 0$ were generated by drawing the noise independently across spatial locations.

assume that neurons with similar phases have stronger noise correlations, because they are likely to share more inputs.

Mathematically, I consider noise autocorrelation functions of the form:

$$C_\xi(\tau_\phi, \tau_x) := \left\langle \xi(\phi, x)\xi(\phi + \tau_\phi, x + \tau_x) \right\rangle_{\phi,x} \tag{6.4}$$

$$= \sigma_\xi^2 G(\tau_\phi; \sigma_\phi) G(\tau_x; \sigma_x) \tag{6.5}$$

where $\sigma_\xi^2 > 0$ is the noise variance and $G(\tau, \sigma) := \exp[-\tau^2/(2\sigma^2)]$ is a Gaussian function with standard deviation $\sigma > 0$. The parameters $0 < \sigma_\phi \ll 2\pi$ and $0 < \sigma_x \ll L$ control the correlation lengths across neurons and across space, respectively (Figure 6.7).

For $\sigma_\phi \to 0$, the noise is uncorrelated across neurons. For increasing values of σ_ϕ, larger and larger groups of neurons receive correlated activity, and correlations are stronger between neurons with nearby phases. Similarly, for $\sigma_x \to 0$, the noise is uncorrelated across spatial locations, and for larger values of σ_x it becomes increasingly smooth in space. Note that I consider only fully-separable autocorrelation functions C_ξ, meaning that correlations across space and neurons are orthogonal. Also note that the input-noise variance σ_ξ^2 is independent of the correlation lengths σ_ϕ and σ_x.

Network activity at the population and single-cell levels

To quantify grid-pattern amplification in the recurrent network, I compute the network activity at two different levels: the single-cell level and the population level (see also Figure 6.5). To this end, I introduce the following notation. I call $h(\phi_0, x)$ the *single-cell* feed-forward input to a neuron with preferred phase ϕ_0, and I call $h(\phi, x_0)$ the *population* feed-forward input to the network for the virtual rat being at location x_0 on the track. Consistently, I call $v(\phi_0, x)$ the single-cell output of neuron ϕ_0, and $v(\phi, x_0)$ the population output at location x_0.

6.4.2. Population-level amplification

In Section 6.B.1, I study the network dynamics in Equation 6.1 at the population level. I show that, if the recurrent connections are not too strong ($M_{\max} < 1$), the population output $v(\phi, x_0)$ converges to a stable fixed point $v_\infty(\phi, x_0)$ for any spatial location x_0 on the track. Importantly, at the steady-state output, the recurrent connections amplify the power of the population input at the first harmonic by a factor

$$A_{\text{pop}} \overset{(6.36)}{=} \frac{1}{(1 - M_{\max})^2} > 1\,. \tag{6.6}$$

Figure 6.8 shows an example of this effect. Because the feed-forward input was assumed to be weakly grid-tuned, a noisy bump of activity emerges at the population

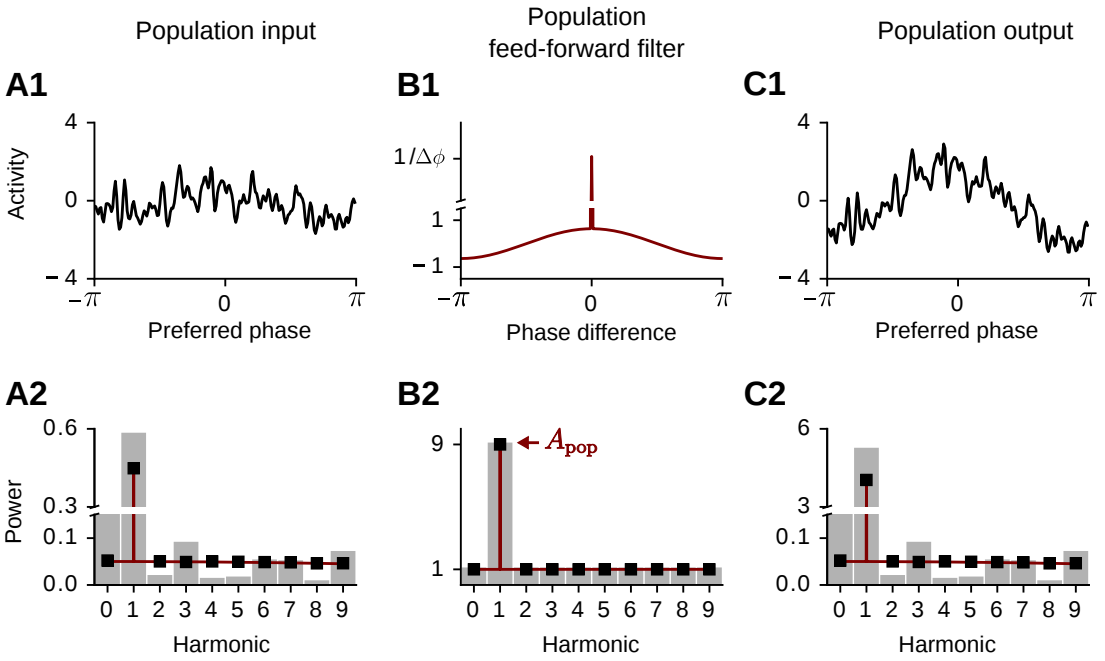

Figure 6.8. | **Recurrent amplification at the population level. A)** Population feed-forward input $h(\phi, x_0)$ to the network for the virtual rat being at location $x_0 = 0$ on the track. A1) Example of the population input $h(\phi, x_0)$ for a single realization of the input noise. Note a noisy bump of activity centered at preferred phase $\phi = 0$. A2) Gray bars: power spectrum of the signal in A1 (only the first ten harmonics are shown). Black squares: average power spectrum estimated from multiple realizations of the input noise. Red line: analytical power spectrum (Section 6.B.2, Equations 6.43, 6.44, and 6.47). **B)** Equivalent population feed-forward filter of the network (B1) and its power spectrum (B2). The red lines in B1 and B2 show the analytical solutions in Equations 6.37 and 6.36, respectively. Gray bars in B2 are obtained by dividing the power spectrum in C2 (gray bars) by the power spectrum in A2 (gray bars). Black squares in B2 are obtained by dividing the average power spectrum in C2 (black squares) by the average power spectrum in A2 (black squares). **C)** Population stead-state output $v_\infty(\phi, x_0)$ of the network for the virtual rat being at location $x_0 = 0$ on the track (see panel A for further details). The analytical power spectrum in C2 (red line) is derived in Section 6.B.2 (Equation 6.48). Note that the first harmonic of the population input is amplified by a factor $A_{\text{pop}} = 9$ at the steady-state output. Parameter values: $M_{\max} = 2/3$, $B = 0.5$, $f = 1\ \text{m}^{-1}$, $L = 5$ m, $\sigma_\xi^2 = 0.4$, $\sigma_x = 0.05$ m, $\sigma_\phi = 0.05$. The phase space is discretized in 200 bins ($\Delta_\phi \approx 0.03$).

input (A1), and the input power is dominated by the first harmonic (A2). At the output (C), the first harmonic of the population input is amplified by a factor $A_{\text{pop}} = 9$ whereas all other harmonics are left unchanged (C2). This effect is quantified by the power of the equivalent *population feed-forward filter* of the network (B), i.e., the ratio of the population output and input power spectra at steady state (Section 6.B.1).

6.4.3. Single-neuron amplification

I now show that the recurrent dynamics also improves the spatial regularity of the input tuning at the single-cell level (Figure 6.9). Because the feed-forward input is weakly grid-tuned with frequency $f = 1\ \text{m}^{-1}$ on a track of length $L = 5$ m (A1), the input power spectrum is dominated by the *tuning harmonic* $Lf = 5$ (A2). Crucially, this harmonic is also amplified at the steady-state output (C). Single-cell amplification is quantified by the power of the equivalent *single-cell feed-forward filter* of the network, i.e., the ratio of the single-cell output and input power spectra at steady state (B, bottom). In the example of Figure 6.9, the feed-forward input is amplified by a factor $A_{\text{cell}} \approx 7.98$ at the tuning harmonic (Equation 6.92) and by a lower factor $A_{\text{noise}} \approx 1.32$ at all other frequencies (Equation 6.83).

Amplification index

To measure the improvement of a single-neuron spatial tuning, I define the *amplification index* Ψ as the ratio between the factors A_{cell} and A_{noise}:

$$\Psi := \frac{A_{\text{cell}}}{A_{\text{noise}}}\,. \tag{6.7}$$

The amplification index Ψ depends on (1) the absolute strength of the recurrent connections; (2) the input-noise correlations across neurons; and (3) the strength of the input grid signal relative to the noise.

In Section 6.B.2 (Equation 6.93), I derive

$$\Psi = \frac{A_{\text{pop}} / A_{\text{noise}}\, \text{SNR}_{\text{in}} + 1}{\text{SNR}_{\text{in}} + 1} \tag{6.8}$$

where SNR_{in} is the *input signal-to-noise ratio*, that is, the ratio of the single-cell power spectra of the input signal g and of the input noise ξ at the tuning harmonic $k = Lf$ (Equation 6.91, Figure 6.9A2). For small values of $\text{SNR}_{\text{in}} \to 0$, the amplification index Ψ approaches 1, meaning that no improvement in spatial tuning is obtained (Figure 6.10A). Conversely, for large values of $\text{SNR}_{\text{in}} \gg 1$, the amplification index Ψ approaches $A_{\text{pop}} / A_{\text{noise}}$, where the factor A_{pop} depends on the connectivity strength M_{max}

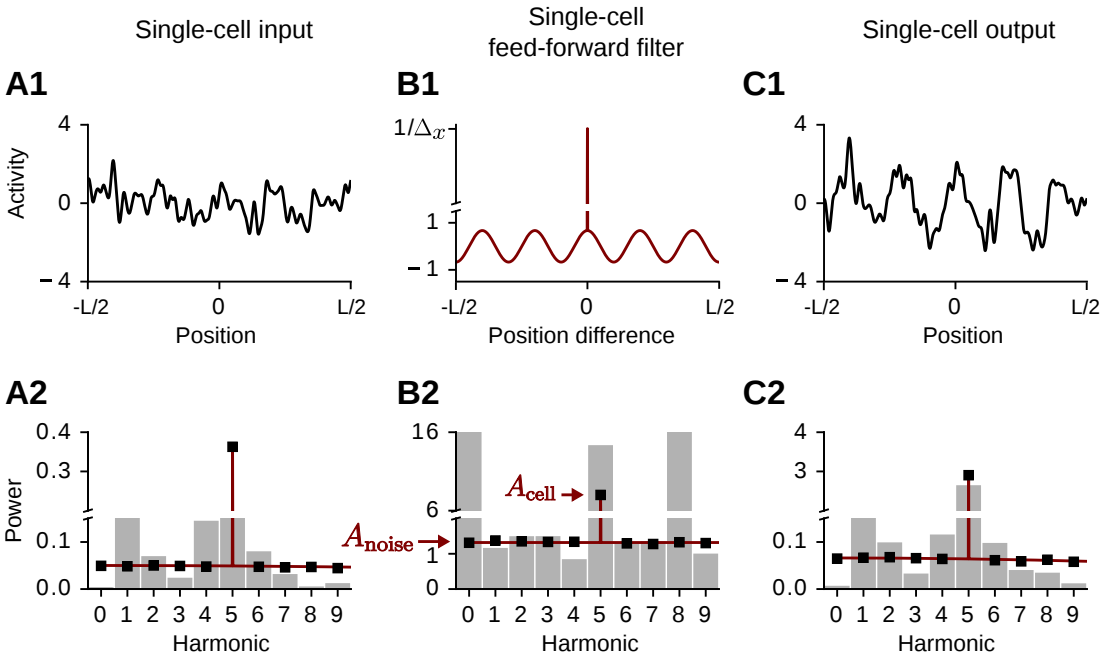

Figure 6.9. | Recurrent amplification at the single-cell level. A) Single-cell feed-forward input $h(\phi_0, x)$ to a cell with preferred phase $\phi_0 = 0$. A1) Example of the single-cell input $h(\phi_0, x)$ for a single realization of the input noise. A2) Gray bars: power spectrum of the trace in A1 (only the first ten harmonics are shown). Black squares: average power spectrum estimated from multiple realizations of the input noise. Red line: analytical power spectrum (Section 6.B.3, Equations 6.54, 6.55, and 6.58). **B)** Equivalent single-cell feed-forward filter of the network (B1) and its power spectrum (B2). The red lines in B1 and B2 depict the analytical solutions in Equations 6.89 and 6.86, respectively. Gray bars in B2 are obtained by dividing the power spectrum in C2 (gray bars) by the power spectrum in A2 (gray bars). Black squares in B2 are obtained by dividing the average power spectrum in C2 (black squares) by the average power spectrum in A2 (black squares). **C)** Single-cell steady-state output $v_\infty(\phi_0, x)$ of a neuron with preferred phase ϕ_0 (see panel A for further details). The analytical power spectrum in panel C2 (red line) is derived in Section 6.B.3 (Equations 6.76, 6.77, and 6.84). The power of the single-cell feed-forward input is amplified by a factor $A_{\text{cell}} \approx 7.98$ at the tuning harmonic $Lf = 5$ (Equation 6.92) and by a factor $A_{\text{noise}} \approx 1.32$ at all other frequencies (Equation 6.83). Parameter values: $B = 0.5$, $f = 1\ \text{m}^{-1}$, $L = 5$ m, $\sigma_\xi^2 = 0.4$, $\sigma_x = 0.05$ m, $\sigma_\phi = 0.05$. Derived quantities: $\text{SNR}_{\text{in}} \approx 6.55$ (Equation 6.91), $\Psi \approx 6.05$ (Equation 6.93). The phase space is discretized in 1000 bins ($\Delta_x = 5 \cdot 10^{-3}$ m).

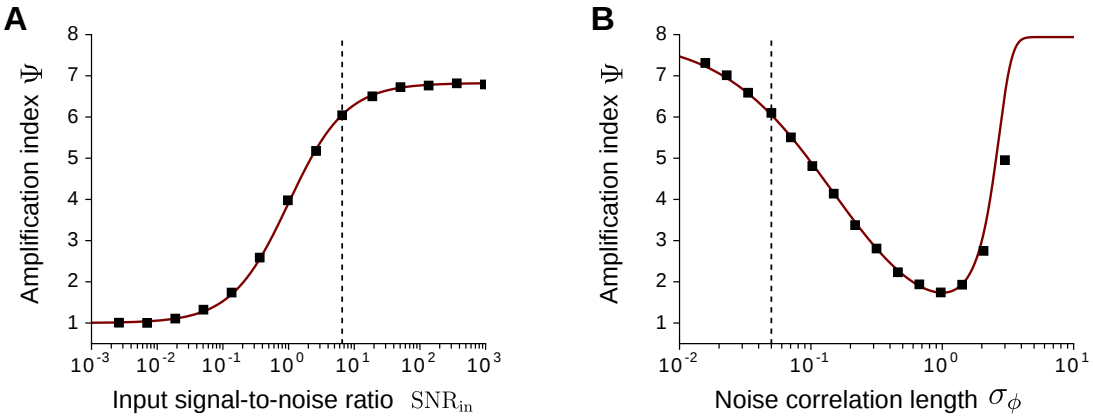

Figure 6.10. | Amplification index. A) Amplification index Ψ as a function of the input signal-to-noise ratio $\mathrm{SNR}_{\mathrm{in}}$. The red line (black squares) depicts analytical (numerical) results (Equation 6.8). Different values of $\mathrm{SNR}_{\mathrm{in}}$ are obtained by varying the signal amplitude B in the range $[0.005, 10]$ (Equation 6.91). Similar results are obtained by varying the noise variance σ_ξ^2 or the input-noise correlation length σ_x. The dashed line indicates the value $\mathrm{SNR}_{\mathrm{in}} \approx 6.55$ for the inputs in Figures 6.8 and 6.9. Parameter values: $M_{\max} = 2/3$, $\sigma_\xi^2 = 0.4$, $\sigma_\phi = 0.05$, $\sigma_x = 0.05$ m. Derived quantities: $A_{\mathrm{pop}} = 9$, $A_{\mathrm{noise}} \approx 1.32$. **B)** Amplification index Ψ as a function of the input-noise correlation length σ_ϕ. The red line (black squares) depicts analytical (numerical) results (Equations 6.8 and 6.9). Parameter values: $M_{\max} = 2/3$, $B = 0.5$, $\sigma_\xi^2 = 0.4$, $\sigma_x = 0.05$. The dashed line indicates the value $\sigma_\phi = 0.05$ for the inputs in Figures 6.8 and 6.9. Derived quantities: $A_{\mathrm{pop}} = 9$, $\mathrm{SNR}_{\mathrm{in}} \approx 6.55$.

(Equation 6.6) and the factor

$$A_{\mathrm{noise}} \overset{(6.83)}{\approx} 1 + (A_{\mathrm{pop}} - 1)\sqrt{\frac{2}{\pi}}\,\sigma_\phi \exp\left(-\frac{\sigma_\phi^2}{2}\right) \tag{6.9}$$

depends on the input-noise correlations across neurons (Equation 6.5).

For uncorrelated input noise ($\sigma_\phi \to 0$), the factor A_{noise} is close to 1, and the amplification is maximal ($\Psi \to A_{\mathrm{pop}}$, Figure 6.10B). For increasing values of σ_ϕ, the factor A_{noise} increases and the index Ψ decreases. The worst case scenario is obtained for $\sigma_\phi = 1$, where A_{noise} its maximal and ψ is minimal (Figure 6.10B). This is because, for $\sigma_\phi = 1$, the first harmonic of the input noise at the population level is largest, and this first harmonic is amplified by the recurrent network (Figure 6.8).

In summary, I showed that structured recurrent connectivity can amplify grid-spatial tuning at the single-cell level, and I quantified this amplification analytically for different input scenarios. The strongest amplification is obtained when the recurrent connections are strong, the input noise is uncorrelated across neurons, and the input signal-to-noise ratio is large.

6.5. Discussion

I studied the dynamics of grid-cell activity in both feed-forward and recurrent networks. First, I showed that grid-cell activity can be inherited via feed-forward projections (Section 6.2.1). Grid-field inheritance required co-modular input grids (i.e., grids with similar spacing and orientation) and structured (i.e., non-random) feed-forward connectivity. I also showed that such a connectivity structure could self-organize through Hebbian synaptic plasticity.

Second, I demonstrated that structured feed-forward projections could even amplify grid-cell activity, i.e., they could improve the regularity of noisy grid patterns (Section 6.2.2). Similarly, grids could be amplified by feedback connections in recurrent circuits (Section 6.3). In this case, structured excitation and unstructured inhibition led to a strong grid tuning of the excitatory cells while keeping the spatial selectivity of the inhibitory neurons low.

Finally, I proposed a minimal mathematical model for the recurrent amplification of grid-cell activity on linear tracks (Section 6.4). In this model, I analytically quantified the amount of amplification a network can achieve for different input-tuning scenarios. In the following, I discuss the main assumptions of this work and I formulate predictions for future experiments.

6.5.1. Origin and inheritance of grid-cell activity

Throughout this study, I assumed feed-forward inputs that were either regular grids (Section 6.2.1) or noisy grids as a function of space (Sections 6.2.2, 6.3, and 6.4). I propose that such an initial tuning could arise via a single-cell learning process based on place-selective inputs and spike-rate adaptation (Chapter 5). Periodic tuning could thus emerge within a sub-population of the entorhinal neurons, and then be inherited (and amplified) across different regions of the MEC (Section 6.2); see also (Tocker et al., 2015) for a similar proposal, and (D'Albis et al., 2015) and (Jaramillo et al., 2014) for related studies on the inheritance of place-cell activity and phase precession.

Because feed-forward projections could amplify grid patterns (Section 6.2.2), I suggest that regions where grids are most abundant, e.g., MEC layer II, are functionally downstream to regions where only fewer grid cells are found, e.g., MEC layers III and V (see Boccara et al., 2010, for a comparison of gridness scores across layers). In this view, grids could originate in layer III via presubicular inputs, or in layer V via hippocampal inputs. Indeed, place-selective firing—which could drive the initial pattern formation process (Chapter 5)—has been observed both in the presubiculum (Taube, 1995b; Cacucci et al., 2004; Boccara et al., 2010) and in the hippocampus (O'Keefe, 1976). Consistently, layer V neurons project to the superficial layers (Van Haeften et al., 2003;

Kloosterman et al., 2003), and layer III neurons contact layer II stellate cells with high rates (Winterer et al., 2017).

Note, however, that layer II pyramidal cells, which exhibit strong grid tuning (Tang et al., 2014; Sun et al., 2015), do not receive feed-forward input neither from layer II stellate cells, nor from layer III pyramidal cells (Winterer et al., 2017). This rises the possibility that layer II pyramids develop grid-like tuning independently, possibly integrating spatially-selective inputs from the adjacent parasubiculum (Cacucci et al., 2004; Hargreaves et al., 2005, 2007; Boccara et al., 2010; Tang et al., 2016). In this scenario, layer II stellates may have the strongest periodic tuning in the MEC, as they integrate inputs from at least two grid-cell populations, i.e., the pyramidal cells in layers II and III (Fuchs et al., 2016; Winterer et al., 2017).

6.5.2. Co-modularity of grid-cell activity

In this study, I investigated the effects of cortical connectivity on grid patterns that were assumed to be co-modular, i.e., they had similar spacing and orientation but distributed spatial phases. Consistently with this assumption, excitatory connectivity sharply decays with somatic distance (Holmgren et al., 2003; Stepanyants and Chklovskii, 2005; Perin et al., 2011), and anatomically nearby grid cells are co-modular (Hafting et al., 2005; Stensola et al., 2012). Considering the anatomical extent of a typical grid-cell module (Stensola et al., 2012), the co-modularity assumption is valid for interlaminar connections at similar dorso-ventral locations, and for intralaminar connections (tangential to the pial surface) within distances of $\sim 100 - 200$ µm (see Figure 4 in Stensola et al., 2012). Notably, pyramidal-cell somata in MEC layer II also cluster in calbindin-positive patches extending $\sim 100 - 150$ µm (Fujimaru and Kosaka, 1996; Kitamura et al., 2014; Ray et al., 2014; Naumann et al., 2016, 2018), which is suggestive of a high recurrent connectivity at this spatial scales.

But what would happen if long-range projections connected different grid-cell modules to a single target? Previous theoretical work has shown that such a scenario leads to sparse spatial selectivity—alike to place fields in the hippocampus (e.g., Solstad et al., 2006; Rolls et al., 2006; Cheng and Frank, 2011). Because place-cell activity is scarce in the entorhinal cortex (Diehl et al., 2017), excitatory connectivity in the MEC is likely to be confined within modules.

6.5.3. Feed-forward versus recurrent amplification

I studied the amplification of grid-cell activity in both feed-forward (Section 6.2.2) and recurrent circuits (Sections 6.3 and 6.4). Here, I discuss differences and similarities between these two network topologies in terms of amplification.

To a first approximation, amplification can be obtained in both feed-forward and

recurrent circuits alike. Accordingly, the activity of a linear (or linearized) recurrent network can always be rewritten in terms of an equivalent feed-forward circuit (Section 6.4); see also (e.g., Ben-Yishai et al., 1995; Carandini and Ringach, 1997; Murphy and Miller, 2009) for similar analytical approaches. Yet these two network topologies produce different results in two key aspects: the temporal dynamics of amplification, and the learning of synaptic weights.

In terms of temporal dynamics, amplification is obtained faster in feed-forward than in recurrent networks. This is because recurrent-network activity transiently reverberates through the feedback connections before relaxing to a steady-state level (Figure 6.4). A steady state is reached in a time that depends on the integration time constant of the network and on the maximal connectivity strength. Notably, because amplification scales with connection strength (Section 6.4, Equations 6.6 and 6.8), a tradeoff exists between the speed of a network's response and the amount of amplification that can be reached—a phenomenon termed 'Hebbian slowing' in the literature (Murphy and Miller, 2009). By contrast, amplification in feed-forward networks is obtained 'at one shot', i.e., its temporal dynamics depends only on the synaptic delays and on the filtering properties of the neurons.

Yet grid-cell activity does not need to be amplified fast. This is because spatially-tuned feed-forward inputs vary slowly in time for typical exploration speeds of the animal, i.e., inputs can be considered approximately constant within the time of a theta cycle ($\sim$100–150 ms). Within this time, a considerable amount of amplification can be obtained in a recurrent circuit (Section 6.3, Figure 6.4).

Note that I consider only the case of structured excitatory connections and random inhibitory weights—a scenario that leads to a broad spatial tuning of the inhibitory neurons (Buetfering et al., 2014). In this case, the network dynamics falls in a regime often called 'Hebbian' or 'normal' amplification (Murphy and Miller, 2009). By contrast, structured inhibitory feedback—if strong enough—could lead to a regime of 'balanced' or 'non-normal' amplification. In this regime, the phenomenon of Hebbian slowing is greatly alleviated, but excitatory and inhibitory neurons also show similar tuning curves—which is not observed in the MEC (Buetfering et al., 2014).

Finally, there are differences between feed-forward and recurrent circuits in terms of learning. A fundamental issue of plastic recurrent networks is to maintain the dynamics of the neuronal activity and the synaptic weights stable. To ensure stability, one needs to avoid a positive-feedback loop in which elevated weights lead to elevated activity that further increases the synaptic weights. Furthermore, because amplification depends on connection strength, the maximal amount of amplification that a recurrent circuit can achieve is limited by stability constraints. By contrast, positive-feedback loops do not emerge in feed-forward circuits and potentially larger amounts of amplification could be reached.

In Section 6.3.1 (Figure 6.3), I simulated the development of the excitatory synaptic weights in a recurrent network. To ensure network stability, I imposed an upper saturation bound $w_{\max} > 0$ to the synaptic weights. The value of $w_{\max}$ was low enough to avoid entering in a positive-feedback loop, but high enough such that recurrent dynamics contributed to the output firing. Importantly, for this setup, I obtained a substantial amount of amplification at the output. Other approaches to guarantee network stability include silencing the recurrent connections during learning (e.g., Stringer et al., 2002b,a; Widloski and Fiete, 2014), or assuming neuronal activation functions that saturate for elevated inputs (e.g., Faugeras et al., 2008; Galtier et al., 2012).

6.5.4. The effects of inhibition on grid-cell tuning

In Section 6.3.2, I simulated the amplification of grid-cell activity in a recurrent network comprising both excitatory and inhibitory neurons (Figure 6.4). I assumed that inhibition was local and unstructured, i.e., a single interneuron received input from (and provided output to) a random set of excitatory cells. This scenario is consistent with experimental data on parvalbumin-positive (PV) interneurons in the MEC, which were found to inhibit grid-cell activity in a phase-independent manner (Buetfering et al., 2014).

Therefore, in my model, inhibition provides a broad baseline shift to the firing rates of the excitatory cells, reducing both in-field and out-of-field spatial firing. This firing-rate reduction also leads to sharper grid fields due to a non-linear transfer in the neuron model (Section 6.6.3). These results are consistent with experimental evidence showing that grid patterns exhibit higher firing rates and broader grid tuning after inactivation of PV interneurons (Buetfering et al., 2014; Miao et al., 2017).

6.5.5. The functional role of recurrent connectivity

The functional role of recurrent connectivity in the cortex has been studied extensively. Recurrent networks are indeed powerful computational systems that can serve a wide range of functions, such as, pattern completion, gain modulation, working memory, and amplification (Rolls et al., 1998; Dayan and Abbott, 2001).

Notably, the recurrent amplification of feed-forward signals has been studied in the visual system thoroughly (e.g., Ben-Yishai et al., 1995; Douglas et al., 1995; Somers et al., 1995; Suarez et al., 1995; Carandini and Ringach, 1997; Murphy and Miller, 2009; Lien and Scanziani, 2013; Ko et al., 2013). In this context, it has been hypothesized that feedback collaterals in V1 could amplify weak directional selectivity provided by thalamic afferents—a conjecture that has been later validated experimentally (Lien and Scanziani, 2013; Ko et al., 2013). Here, I propose that in a similar fashion recur-

rent connectivity in the MEC could amplify feed-forward signals with weak spatial periodicity.

Recurrent dynamics is also at the core of CAN models of grid-cell activity (e.g., McNaughton et al., 2006; Fuhs and Touretzky, 2006; Guanella et al., 2007; Burak and Fiete, 2009, see also Section 4.1). In these models, however, feedback projections are thought to generate spatial selectivity, rather than 'simply' amplify feed-forward tuning. In the following, I propose a framework in which CAN and amplification models can be partly reconciled (see also Section 6.3.3).

CANs can generate grids from completely aperiodic inputs by using structured recurrent connectivity and ideothetic self-motion inputs. In this case, the periodicity of the spatial patterns directly follows from the topology of the recurrent connections (see also Section 4.1.3 for more details). But how could such a recurrent network topology arise in the first place?

Widloski and Fiete (2014) proposed that structured recurrent connectivity could be learned from spatially-tuned (but aperiodic) feed-forward inputs and ideothetic (velocity-modulated) signals. In this case, grid patterns in the mature network emerge from the integration of velocity inputs alone. Therefore, the grid scale depends on the integration speed of the network, which in turn is set by the temporal asymmetry of the learning rules during development (see also Section 4.1.4). Yet to obtain stable single-cell grids a frequent calibration with spatially-tuned inputs is required (Burak and Fiete, 2009, see also Section 4.1.5).

Therefore, spatial inputs may be required 1) for learning the recurrent connections during development, and 2) for anchoring the network activity during exploration. If the same inputs were used both for learning and for anchoring, spatially-selective feed-forward inputs and velocity-modulated recurrent signals shall be always in agreement during exploration. This, however, was not the case in the model by Widloski and Fiete (2014). Alternatively, different sets of feed-forward inputs could be used for learning and anchoring. In this scenario, however, new feed-forward connections must be learned after the development of the recurrent network (see, e.g., Hardcastle et al., 2015).

Here, I propose an alternative scenario. During development, weakly-periodic grid patterns emerge from a single-cell process driven by spatially-tuned inputs and spike-rate adaptation (Chapter 5). The scale of the patterns is given by the temporal dynamics of adaptation, the width of the input tuning, and the average speed of the rat during learning. Shortly after this initial learning phase, activity-dependent synaptic plasticity connects grids with similar phases generating a periodic recurrent connectivity similar to the one assumed by CAN models (Section 6.3). This recurrent connectivity could serve two functions: 1) it could generate sharper grids when the feed-forward input is weakly grid tuned; 2) it could maintain grid-like activity for short time stretches in

which the feed-forward tuning is temporarily lost.

The latter function—which partially reconciles CAN and amplification models—has not been explicitly modeled in this thesis. The idea is the following. After the recurrent connections between grid cells have developed, associations with velocity-modulated inputs could be also learned. These connections could sustain grid-cell activity based on self-motion information alone—though only for brief periods of time. Importantly, the integration speed of the network now matches the tuning of the feed-forward inputs, i.e., the scale of the input grids. Therefore, in this scenario, the associations with the self-motion inputs are learned from the spatially-tuned feed-forward signals and not vice-versa.

I propose that this theory could explain the emergence, amplification, and maintenance of grid-cell activity in a simpler conceptual framework as the one suggested by standard CAN models. Indeed, the same spatial inputs could be used both for circuit development and for anchoring during exploration. Additionally, the scale of the resulting grids would depend only on simple biophysical variables (such as the time scale of adaptation) rather than on the details of the plasticity rules used during development.

6.5.6. Model predictions and conclusions

A central result of the present work is that grid-cell activity can be inherited and amplified via feed-forward projections. This hypothesis could be tested in the connection from the pyramidal to the stellate cells in MEC layer II, which is particularly strong and not reciprocated (13.5% connection probability, Winterer et al., 2017). I thus predict that inactivating layer II pyramidal cells shall result in a reduced grid tuning of layer II stellate cells. A similar effect, though perhaps weaker, could be observed by inactivating pyramidal cells in layer III (9% connection probability to layer II stellates, Winterer et al., 2017).

Another finding of the present study is that Hebbian synaptic plasticity can generate structured (i.e., phase specific) recurrent connectivity between grid cells after a weak periodic tuning has emerged at the feed-forward input. Indirect experimental evidence suggests that such a structured connectivity is indeed present in mature grid-cell networks (Yoon et al., 2013; Dunn et al., 2015; Tocker et al., 2015). However, whether this structured connectivity emerges through activity-dependent processes during development is still unclear.

A similar question has been addressed in visual cortex. To study the emergence of functionally-organized cortical circuits, Ko et al. (2013) characterized the tuning curves of nearby cells in V1 through two-photon optical imaging *in vivo*, and subsequently tested for connectivity patterns between the same cells using whole-cell recordings *in*

vitro. By repeating the experiment for animals at different stages of developmental maturation and visual experience, the authors were able to assess the development of synaptic connectivity in visual cortex. They found that, with experience, local connectivity reorganized extensively and matched the activity profiles characterized *in vivo* (Ko et al., 2013). I suggest that a similar experimental framework could shed light on the experience-dependent maturation of grid-cell microcircuits in the MEC.

In Section 6.5.5, I discussed two alternative theories by which grid-cell activity could develop. In the CAN model by Widloski and Fiete (2014), grid-cell activity develops via spatially-selective signals and velocity-modulated inputs (e.g., speed-modulated head-direction cells). By contrast, in the theory presented here (Section 6.5.5), grid-pattern formation initially requires only spatially-tuned inputs and spike-rate adaptation, but not velocity inputs. Therefore, I predict that inactivating speed or head-direction cells during development shall spare a weak periodic tuning of the grid cells in the adult animal.

Finally, I showed that local recurrent circuits can amplify grid-cell patterns (Section 6.3). This hypothesis could be experimentally tested by silencing the output activity of a recurrently-connected grid-cell population while recording excitatory post-synaptic currents at the feed-forward input. Such an approach has been proven successful in demonstrating the recurrent amplification of thalamic signals in visual cortex (Lien and Scanziani, 2013). However, isolating feed-forward and recurrent excitation in sub-populations of entorhinal neurons may be more difficult than in visual pathways.

In summary, I showed that entorhinal grid-cell activity can be inherited and amplified by excitatory networks, and that these networks can self-organize in an unsupervised manner. Finally, I proposed a coherent framework in which the emergence, amplification, and maintenance of grid-cell activity is explained at the level of functional microcircuits of the MEC.

6.6. Materials and Methods

6.6.1. Model of neural activity in the feed-forward network

I model a feed-forward network of N_{in} excitatory inputs with firing rates $\{r_j^{\text{in}} : j = 1, 2, \ldots, N_{\text{in}}\}$ projecting to N_{out} excitatory outputs with firing rates $\{r_i^{\text{out}} : i = 1, 2, \ldots, N_{\text{out}}\}$. For each position $\mathbf{x} = [x_1, x_2]$ in the environment, the output firing rates are

$$r_i^{\text{out}}(\mathbf{x}) := \left[\sum_{j=1}^{N_{\text{in}}} w_{ij}\, r_j^{\text{in}}(\mathbf{x}) - r_0 \right]_+ \tag{6.10}$$

where $w_{ij} \geq 0$ is the synaptic weight from input neuron j to output neuron i, the parameter $r_0 > 0$ is a spatially-homogeneous inhibitory rate, and the function $[z]_+ = z$ if $z > 0; = 0$ if $z \leq 0$ is a static non-linearity.

The synaptic weights are initially random, and they are plastic according to a Hebbian learning rule (Section 6.6.2). The spatial tuning of the input rates $\{r_j\}$ is described in Section 6.6.5.

Parameter values: $N_{\text{in}} = 400$, $N_{\text{out}} = 200$, $r_0 = 1$ spike/s.

6.6.2. Model of synaptic plasticity in the feed-forward network

The excitatory synaptic weights w_{ij} are plastic according to the Hebbian rule

$$\frac{\mathrm{d}w_{ij}}{\mathrm{d}t} := \eta (r_j^{\text{in}} - \gamma)\, r_i^{\text{out}} \quad \text{with} \quad 0 \leq w_{ij} \leq w_{\max} \tag{6.11}$$

where $\eta \ll 1$ is a small learning rate and $\gamma > 0$ sets the threshold between synaptic potentiation and depression. At the initial condition, a random subset of $N^{\text{up}} < N^{\text{in}}$ synaptic weights are set to the upper bound $w_{\max} > 0$, whereas all other weights are set to zero.

The synaptic weights are updated every 30 ms while a virtual rat explores a square enclosure for 400 minutes (Section 6.6.6). The weight constraints $0 \leq w_{ij} \leq w_{\max}$ are enforced at each time step of the simulation. I set $w_{\max} \approx r_0/(N_{\text{up}} r_{\text{av}})$ where $r_{\text{av}} > 0$ is the average input firing-rate in the environment (Section 6.6.5). This choice ensures that the number of potentiated synapses remains roughly constant during learning (Section 6.A.1).

Parameter values: $\eta = 2 \cdot 10^{-5}$ and $\gamma = 8$ spikes/s for simulations with regular grids (Figure 6.1) and jittered firing-field centers (Figure 6.2A); $\eta = 5 \cdot 10^{-5}$ and $\gamma = 3.8$ spikes/s for simulations with grids distorted by additive noise (Figure 6.2B). In all simulations of the feed-forward network, I set $N^{\text{up}} = 15$ and $w_{\max} = 0.02$.

6.6.3. Model of neural activity in the recurrent network

I model the activity of a recurrent network comprising a population of N_{E} excitatory neurons with rates $\{r_i^{\text{E}} : i = 1, 2, \ldots, N_{\text{E}}\}$ and a population of N_{I} inhibitory neurons with rates $\{r_i^{\text{I}} : i = 1, 2, \ldots, N_{\text{I}}\}$. The two populations are recurrently connected locally, and they are mutually coupled via feed-forward projections (Figure 6.3A). The external drive to the network is provided by spatially-tuned feed-forward inputs $\{h_i : i = 1, 2, \ldots, N_{\text{E}}\}$ that project exclusively to the excitatory population (Section 6.6.5).

The network activity follows a classical firing-rate description (Wilson and Cowan,

1972, 1973). The firing rates of the excitatory neurons are

$$\tau \frac{\mathrm{d}r_i^{\mathrm{E}}(t)}{\mathrm{d}t} := -r_i^{\mathrm{E}}(t) + \left[h_i(t) + \sum_{j=1}^{N_E} w_{ij}^{\mathrm{EE}} \, r_j^{\mathrm{E}}(t) - \sum_{j=1}^{N_I} w_{ij}^{\mathrm{EI}} \, r_j^{\mathrm{I}}(t) \right]_+ \tag{6.12}$$

where $\tau > 0$ is the integration time constant of the network, $h_i \geq 0$ is the total feed-forward input to neuron i, $w_{ij}^{\mathrm{EE}} \geq 0$ is the synaptic weight from excitatory neuron j to excitatory neuron i, and $w_{ij}^{\mathrm{EI}} \geq 0$ is the synaptic weight from inhibitory neuron j to excitatory neuron i, and the function $[z]_+$ is a static non-linearity. Similarly, the firing rates of the inhibitory neurons are

$$\tau \frac{\mathrm{d}r_i^{\mathrm{I}}(t)}{\mathrm{d}t} := -r_i^{\mathrm{I}}(t) + \left[\sum_{j=1}^{N_E} w_{ij}^{\mathrm{IE}} \, r_j^{\mathrm{E}}(t) - \sum_{j=1}^{N_I} w_{ij}^{\mathrm{II}} \, r_j^{\mathrm{I}}(t) \right]_+ \tag{6.13}$$

where $w_{ij}^{\mathrm{IE}} \geq 0$ is the synaptic weight from excitatory neuron j to inhibitory neuron i and $w_{ij}^{\mathrm{II}} \geq 0$ is the synaptic weight from inhibitory neuron j to inhibitory neuron i.

The synaptic connectivity is sparse and initially random. That is, each neuron in the network receives input from a random subset of $N_{\mathrm{up}}^{\mathrm{E}} < N^{\mathrm{E}}$ excitatory neurons, and from a random subset of $N_{\mathrm{up}}^{\mathrm{I}} < N^{\mathrm{I}}$ inhibitory neurons. For each connection, the corresponding synaptic weight is set to $w_{\max} > 0$. The excitatory synaptic weights w_{ij}^{EE} are plastic according to a Hebbian learning rule (Section 6.6.4).

For the results in Figure 6.3B (learning of the recurrent weights), the network activity is simulated while a virtual rat performs a random walk in the environment for 500 s (Section 6.6.6). For the results in Figures 6.4 and 6.5, the network activity is simulated for each spatial location independently for 200 ms. In both cases, the network activity (Equations 6.12 and 6.13) is updated with the forward Euler method with integration time step $\Delta t = 2$ ms.

Parameter values: $N_{\mathrm{E}} = 900$, $N_{\mathrm{I}} = 225$, $N_{\mathrm{up}}^{\mathrm{E}} = 55$, $N_{\mathrm{up}}^{\mathrm{I}} = 70$, $\tau = 10$ ms, $w_{\max} = 0.015$.

6.6.4. Model synaptic plasticity in the recurrent network

The excitatory synaptic weights w_{ij}^{EE} are plastic according to the Hebbian rule:

$$\frac{\mathrm{d}w_{ij}^{\mathrm{EE}}}{\mathrm{d}t} := \eta (r_i^{\mathrm{E}} - r_{\mathrm{av}})(r_j^{\mathrm{E}} - r_{\mathrm{av}}) \quad \text{with} \quad 0 \leq w_{ij}^{\mathrm{EE}} \leq w_{\max} \tag{6.14}$$

where $\eta \ll 1$ is a small learning rate, $r_i^{\mathrm{E}} \geq 0$ is the firing rate of excitatory neuron i, and $r_{\mathrm{av}} > 0$ is the average input firing rate in the environment (Section 6.6.5). Additionally, at each time step, the weights are additively normalized to ensure that the total input

and output weight of each neuron is kept constant during learning, i.e.,

$$\sum_{j=1}^{N_E} w_{ij}^{EE} = \sum_{i=1}^{N_E} w_{ij}^{EE} = N_{up}^{E} w_{max} \tag{6.15}$$

where $N_{up}^{E} < N_E$ is the number of input excitatory weights set to the upper bound w_{max} at the initial condition (Section 6.6.3).

During learning, the synaptic weights are updated every 30 ms while a virtual rat explores a square enclosure for 500 s (Section 6.6.6). The weight constraints $0 \leq w_{ij}^{EE} \leq w_{max}$ are enforced at each time step of the simulation.

Parameter values: $\eta = 0.01$, $w_{max} = 0.015$, $N_E = 900$, $N_E^{up} = 55$, $r_{av} = 3$ spikes/s.

6.6.5. Model of input spatial tuning

Here, I describe the spatial tuning of the inputs to both the feed-forward and the recurrent network (Sections 6.2 and 6.3). Three input scenarios are considered: 1) Grid-tuned inputs; 2) Weakly grid-tuned inputs; and 3) Spatially-aperiodic inputs.

Grid-tuned inputs are used to illustrate the inheritance of grid-cell activity in the feed-forward network (Figure 6.1). Weakly grid-tuned inputs are used to illustrate the amplification of grid-cell activity in both feed-forward (Figure 6.2) and recurrent networks (Figures 6.3, 6.4, and 6.5A). Spatially-aperiodic inputs are used to illustrate the relation between amplification and CAN models in the recurrent network (Figure 6.5B).

Grid-tuned inputs

Grid-tuned inputs are generated by summing three planar waves with wave vectors that are 60 degrees apart:

$$g_j(\mathbf{x}) = A_i \left[\sum_{n=1}^{3} \cos\left(\mathbf{k}_{jn} \cdot (\mathbf{x} + \boldsymbol{\varphi}_j) \right) \right]_+ \quad \text{with} \quad \mathbf{k}_{jn} := \frac{4\pi}{T_j \sqrt{3}} \begin{bmatrix} \cos(\frac{n\pi}{3} + \alpha_j) \\ \sin(\frac{n\pi}{3} + \alpha_j) \end{bmatrix} \tag{6.16}$$

where $\mathbf{x}$ is the position of the virtual rat in the environment, $[z]_+$ is a static non-linearity, $A_j > 0$ scales the peak firing rate, T_j is the grid spacing, α_j is the grid orientation, and $\boldsymbol{\varphi}_j$ is the grid spatial phase. Note that all planar waves have spatial period $T_j\sqrt{3}/2$, which generates a grid spacing T_j.

The scale factors $\{A_j\}$ are chosen such that all grids are normalized to the same average rate in the environment, i.e., $(1/L^2) \sum_{\mathbf{x}} g_j(\mathbf{x}) = r_{av} \ \forall j$. The grid spacings $\{T_j\}$ and the grid orientations $\{\alpha_j\}$ are normally distributed with means μ_T and μ_α and standard deviations σ_T and σ_α, respectively. The grid phases $\{\boldsymbol{\varphi}_j\}$ are sampled evenly in phase space.

Parameter values: $r_{av} = 3$ spikes/s, $\mu_T = 0.5$ m, $\sigma_T = 0.05$ m, $\mu_\alpha = 0$, $\sigma_\alpha = 0.1$.

Weakly grid-tuned inputs

Two types of weakly-grid tuned inputs are considered: grids with jittered firing-field centers (Figure 6.2A) and grids distorted by additive noise (Figures 6.2B, 6.3, 6.4, and 6.5A).

Grids with jittered firing-field centers

Grids with jittered firing-field centers are generated as follows. First, co-modular regular grids are generated (Equation 6.16) and their firing-field centers numerically detected. Second, each firing-field center is displaced by means of a two-dimensional shift vector whose components are normally distributed with zero mean and standard-deviation σ_{jitter}. Third, a spatial map is constructed by superimposing a two-dimensional Gaussian function for each jittered firing-field center:

$$g_j^{\text{jitter}}(\mathbf{x}) := A_j^{\text{jitter}} \left[\sum_k \exp\left(-\frac{(\mathbf{x} - \mathbf{c}_{jk})^2}{2\sigma_{\text{field}}^2} \right) - 0.5 \right]_+ \tag{6.17}$$

where $\{\mathbf{c}_{jk}\}$ are the jittered firing-field centers of input grid j and $\sigma_{\text{field}} > 0$ sets the width of the firing fields. To mimic the firing-rate profiles of the original grids (Equation 6.16), a baseline shift and a static non-linearity are introduced in Equation 6.17. Finally, the scale factors $\{A_j^{\text{jitter}}\}$ are chosen such that all grids are normalized to the same average firing rate r_{av}.

Parameter values: $\sigma_{\text{jitter}} = 0.05$ m, $\sigma_{\text{field}} = 0.135$, $r_{\text{av}} = 3$ spikes/s.

Grids distorted by additive noise

Grids distorted by additive noise are generated by superimposing random spatial maps to regular grids:

$$g_j^{\text{noise}}(\mathbf{x}) := A_j^{\text{noise}} \Big[\beta\, g_j(\mathbf{x}) + (1 - \beta)\, \xi_j(\mathbf{x}) \Big]_+ \tag{6.18}$$

where $g_j(\mathbf{x})$ is a regular grid (Equation 6.16), and $\xi_j(\mathbf{x})$ is a random spatial map. The parameter $0 \leq \beta \leq 1$ controls the strength of the grid signal relative to the noise. The noise is uncorrelated across neurons but correlated (i.e., smooth) in space. Each random map $\xi_j(\mathbf{x})$ is generated independently by filtering 2-dimensional white Gaussian noise with a circularly-symmetric Gaussian filter with standard deviation σ_{noise}. The resulting random maps are then normalized to match the mean and the variance of the grid signal. Note that the maps $\xi_j(\mathbf{x})$ assume both positive and negative values. Finally, the scale factors $\{A_j^{\text{jitter}}\}$ are chosen such that all grids share the same average firing-rate r_{av} in the environment.

Parameter values: $\beta = 0.35$ m, $\sigma_{\text{noise}} = 0.08$ m, $r_{\text{av}} = 3$ spikes/s.

Spatially aperiodic inputs

Spatially-aperiodic inputs are generated with Equation 6.18 setting $\beta = 0$.

Parameter values: $\beta = 0$ m, $\sigma_{\text{noise}} = 0.08$ m, $r_{\text{av}} = 3$ spikes/s.

6.6.6. Model of spatial exploration

The virtual rat explores a square enclosure of side-length L with an isotropic random walk at constant speed $v > 0$. Virtual-rat trajectories $\mathbf{x}_t$ are sampled from the stochastic process

$$\frac{\mathrm{d}\mathbf{X}_t}{\mathrm{d}t} := v\,[\cos(\theta_t), \sin(\theta_t)] \quad \text{with} \quad \theta_t = \sigma_\theta \mathcal{W}_t\,, \tag{6.19}$$

where the angle θ_t sets the direction of motion and $\mathcal{W}_t$ is a standard Wiener process. The parameter σ_θ controls the tortuosity of the trajectory. At the boundaries of the environment, the component of the movement direction perpendicular to the boundary is inverted. Spatial locations in the arena are discretized in 50^2 square bins. The random walk is updated every 30 ms of the simulation.

Parameter values: $L = 1.5$ m, $v = 0.25$ m/s, $\sigma_\theta = 0.7$.

6.7. Chapter Summary

In this chapter, I studied how feed-forward and recurrent circuits could affect grid-cell activity in the entorhinal cortex. First, I showed that grids could be inherited across neuronal populations via structured feed-forward projections. Then, I showed that both feed-forward and recurrent circuits could serve to amplify grid-cell patterns, i.e., they could improve the spatial regularity of noisy inputs. Finally, I outlined a mechanism by means of which recurrent connectivity and self-motion inputs could maintain grid-like firing even when the feed-forward tuning is temporarily lost. In the next chapter, I will summarize the main results and predictions of this thesis.

6.A. Analytical results on the feed-forward network model

In this appendix, I present analytical derivations on the feed-forward network model in Section 6.2.

6.A.1. Weight normalization in the feed-forward network

Here, I show that the plasticity rule for the synaptic weights w_{ij} (Section 6.6.2, Equation 6.11)

$$\frac{\mathrm{d}w_{ij}}{\mathrm{d}t} := \eta(r_j^{\mathrm{in}} - \gamma)\, r_i^{\mathrm{out}} \quad \text{with} \quad 0 \leq w_{ij} \leq w_{\max} \tag{6.20}$$

leads to an implicit normalization of the average input weight

$$w_{\mathrm{av}} := \frac{1}{N_{\mathrm{in}}} \sum_{j=1}^{N_{\mathrm{in}}} w_{ij} \quad \forall i\,. \tag{6.21}$$

Recall that $\eta \ll 1$ is a small learning rate, and that the parameter $\gamma > 0$ sets the threshold between potentiation and depression. The input rates r_j^{in} are spatially tuned and normalized to the same average rate r_{av} (Section 6.6.5). The output rates are given by (Equation 6.10)

$$r_i^{\mathrm{out}} := \left[\sum_{k=1}^{N_{\mathrm{in}}} w_{ik}\, r_k^{\mathrm{in}} - r_0 \right]_+ \tag{6.22}$$

where $r_0 > 0$ is a constant inhibitory rate.

I now study the linear dynamics of the synaptic weights that is obtained by ignoring the saturation boundaries in Equation 6.20 and the static non-linearity in Equation 6.22. I plug Equation 6.22 into Equation 6.20, and I average over a time window that separates the time scale of learning from the time scale of the input changes:

$$\eta^{-1}\frac{\mathrm{d}\overline{w}_{ij}}{\mathrm{d}t} = \sum_{k=1}^{N_{\mathrm{in}}} \underbrace{\overline{r_j^{\mathrm{in}} r_k^{\mathrm{in}}}}_{=:\, C_{jk}}\, \overline{w}_{ik} - \gamma r_{\mathrm{av}} \underbrace{\sum_{k=1}^{N_{\mathrm{in}}} \overline{w}_{ik}}_{=\, N_{\mathrm{in}}\overline{w}_{\mathrm{av}}} + r_0(\gamma - r_{\mathrm{av}}) \tag{6.23}$$

where the bar denotes temporal averaging, C_{ij} is the input correlation matrix, and I assume $\overline{r}_j^{\mathrm{in}} \approx r_{\mathrm{av}}$ for all inputs j. I now average both sides of Equation 6.23 across the synaptic index j:

$$\begin{aligned}
\eta^{-1}\frac{\mathrm{d}\overline{w}_{\mathrm{av}}}{\mathrm{d}t} &= \frac{1}{N_{\mathrm{in}}} \sum_{k=1}^{N_{\mathrm{in}}} \underbrace{\sum_{j=1}^{N_{\mathrm{in}}} C_{jk}}_{\approx\, N_{\mathrm{in}} r_{\mathrm{av}}^2}\, \overline{w}_{ik} - N_{\mathrm{in}}\gamma r_{\mathrm{av}}\overline{w}_{\mathrm{av}} + r_0(\gamma - r_{\mathrm{av}}) && (6.24)\\
&= N_{\mathrm{in}} r_{\mathrm{av}}^2 \overline{w}_{\mathrm{av}} - N_{\mathrm{in}}\gamma r_{\mathrm{av}}\overline{w}_{\mathrm{av}} + r_0(\gamma - r_{\mathrm{av}}) && (6.25)\\
&= -N_{\mathrm{in}} r_{\mathrm{av}}(\gamma - r_{\mathrm{av}})\overline{w}_{\mathrm{av}} + r_0(\gamma - r_{\mathrm{av}})\,. && (6.26)
\end{aligned}$$

Equation 6.26 shows that, for $\gamma > r^{\mathrm{av}}$, the average input weight $\overline{w}_{\mathrm{av}}$ exponentially

approaches a steady-state value w^*_{av} with time constant τ_{av}, where

$$w^*_{\text{av}} = \frac{r_0}{N_{\text{in}} r_{\text{av}}} \quad \text{and} \quad \tau_{\text{av}} = \frac{1}{\eta r_{\text{av}} N_{\text{in}} (\gamma - r_{\text{av}})} . \tag{6.27}$$

Before learning, a random subset of N_{up} synaptic weights is potentiated, i.e, it is set to the upper saturation boundary $w_{\max} > 0$, whereas all other weights are set to zero. To ensure that the fraction of potentiated weights remains roughly constant during learning, I impose

$$w^*_{\text{av}} = \frac{N_{\text{up}}\, w_{\max}}{N_{\text{in}}} , \tag{6.28}$$

and from Equations 6.27 and 6.27, I obtain

$$w_{\max} = \frac{r_0}{r_{\text{av}}\, N_{\text{up}}} . \tag{6.29}$$

6.B. Analytical results on grid-pattern amplification

In this appendix, I present the analytical derivations of the reduced amplification model described in Section 6.4.

6.B.1. Population-level amplification

Here, I study the activity of a recurrent network with weakly grid-tuned feed-forward inputs as a virtual rat runs on a linear track (Section 6.4.1). All input grids have the same spatial frequency f, but different spatial phases ϕ. The output activity v of a neuron with input phase ϕ is (Equation 6.1)

$$\tau \frac{\mathrm{d}}{\mathrm{d}t} v(\phi, x) := -v(\phi, x) + h(\phi, x) + \int_{-\pi}^{\pi} \mathrm{d}\phi'\, M(\phi - \phi') v(\phi', x) \tag{6.30}$$

where x denotes a position on the track, h is the total feed-forward input, the function M defines the recurrent connectivity, and $\tau > 0$ sets the integration time constant of the network.

First, I study the network activity at the population level (Section 6.4.2). To this end, I apply the complex Fourier series transformation

$$\widetilde{f}_k := \int_{-\pi}^{\pi} \mathrm{d}\phi\, f(\phi) \exp(-jk\phi) \quad \Leftrightarrow \quad f(\phi) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \widetilde{f}_k \exp(jk\phi) \tag{6.31}$$

at both sides of Equation 6.30:

$$\tau \frac{\mathrm{d}}{\mathrm{d}t} \widetilde{v}_k(x) = -(1 - \widetilde{M}_k) \widetilde{v}_k(x) + \widetilde{h}_k(x) \tag{6.32}$$

where $\widetilde{v}_k \in \mathbb{C}$, $k \in \mathbb{Z}$, and $j = \sqrt{-1}$ is the imaginary unit. Note that the Fourier coefficients $\widetilde{M}_k$ are the eigenvalues of the dynamical system in Equation 6.30, and the corresponding eigenfunctions are the elements of the Fourier basis $e_k(\phi) = \exp(jk\phi)$.

Equation 6.32 shows that for $\widetilde{M}_k < 1 \ \forall k$, the output activity v converges to a stable fixed point v_∞ with Fourier coefficients

$$\widetilde{v}_k^\infty(x) = \frac{1}{1-\widetilde{M}_k}\widetilde{h}_k(x) = \widetilde{F}_k^{\text{pop}}\widetilde{h}_k(x) \tag{6.33}$$

where the coefficients

$$\widetilde{F}_k^{\text{pop}} := \frac{1}{1-\widetilde{M}_k} \tag{6.34}$$

define the equivalent *population feed-forward filter* of the network. Using the definition of the connectivity function $M(\phi) := 1/\pi\, M_{\max}\cos(\phi)$ (Equation 6.2), I derive

$$\widetilde{M}_k = \begin{cases} M_{\max} & |k| = 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \widetilde{F}_k^{\text{pop}} = \begin{cases} \dfrac{1}{1-M_{\max}} & |k| = 1 \\ 1 & \text{otherwise} \end{cases} \tag{6.35}$$

where the parameter $M_{\max} > 0$ sets the maximal connection strength in the network. Finally, from Equation 6.35, I compute the power spectrum

$$|\widetilde{F}_k^{\text{pop}}|^2 = \begin{cases} A_{\text{pop}} & |k| = 1 \\ 1 & \text{otherwise} \end{cases} \quad \text{with} \quad A_{\text{pop}} := \frac{1}{(1-M_{\max})^2} \tag{6.36}$$

of the equivalent population feed-forward filter

$$F^{\text{pop}}(\phi) \overset{(6.31)}{=} \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\widetilde{F}_k^{\text{pop}}\exp(jk\phi) \overset{(6.36)}{=} \frac{\sqrt{A_{\text{pop}}}-1}{\pi}\cos(\phi) + \delta(\phi) \tag{6.37}$$

where δ is the Dirac's delta.

6.B.2. Population-level power spectra

Here, I compute the population-level power spectra of the feed-forward input h and of the steady-state output v_∞ (Section 6.4.2 and Figure 6.8).

Population-level input power spectrum

I define the population-level input power spectrum

$$S_{\text{pop}}^h(k) := \int_{-\pi}^{\pi} \mathrm{d}\tau\ C_{\text{pop}}^h(\tau)\exp\left(-jk\tau\right) \quad \text{with} \quad k \in \mathbb{Z} \tag{6.38}$$

where

$$C^h_{\text{pop}}(\tau) := \left\langle h(\phi,x)h(\phi+\tau,x)\right\rangle_\phi \tag{6.39}$$

$$\overset{(6.3)}{=} \underbrace{\left\langle g(\phi,x)g(\phi+\tau,x)\right\rangle_\phi}_{=: C^g_{\text{pop}}(\tau)} + \underbrace{\left\langle \xi(\phi,x)\xi(\phi+\tau,x)\right\rangle_\phi}_{=: C^\xi_{\text{pop}}(\tau)} \quad \forall x \tag{6.40}$$

is the population-level autocorrelation of the feed-forward input h (Equation 6.3). In Equations 6.39 and 6.40, g is the input signal, ξ is the input noise, and the angular brackets denote statistical expectation. Note that, because signal and noise are generated independently, the cross terms of the product in Equation 6.39 vanish. In Equation 6.40, the function

$$C^g_{\text{pop}}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{d}\phi\, g(\phi,x)g(\phi+\tau,x) \overset{(6.3)}{=} \frac{B^2}{2}\cos(\tau) \tag{6.41}$$

is the population-level autocorrelation of the input signal g. Similarly, the function

$$C^\xi_{\text{pop}}(\tau) = C_\xi(\tau,0) \overset{(6.5)}{=} \sigma^2_\xi G(\tau;\sigma_\phi) \quad \text{with} \quad G(\tau,\sigma) := \exp\left(-\frac{\tau^2}{2\sigma^2}\right) \tag{6.42}$$

is the population-level autocorrelation of the input noise ξ (Section 6.4.1).

From Equations 6.38 and 6.40, I find that the population-level power spectrum S^h_{pop} of the feed-forward input h is the sum of the population-level power spectra S^g_{pop} and S^ξ_{pop} of the input signal g and of the input noise ξ, respectively. That is,

$$S^h_{\text{pop}}(k) = S^g_{\text{pop}}(k) + S^\xi_{\text{pop}}(k) \tag{6.43}$$

where

$$S^g_{\text{pop}}(k) := \int_{-\pi}^{\pi} \mathrm{d}\tau\, C^g_{\text{pop}}(\tau)\exp\left(-jk\tau\right) \overset{(6.41)}{=} \begin{cases} \dfrac{B^2\pi}{2} & |k| = 1 \\ 0 & \text{otherwise} \end{cases} \tag{6.44}$$

and

$$S^\xi_{\text{pop}}(k) := \int_{-\pi}^{\pi} \mathrm{d}\tau\, C^\xi_{\text{pop}}(\tau)\exp\left(-jk\tau\right) \tag{6.45}$$

$$\overset{(6.42)}{\approx} \sigma^2_\xi \int_{-\infty}^{\infty} \mathrm{d}\tau\, G(\tau;\sigma_\phi)\exp\left(-jk\tau\right) \tag{6.46}$$

$$= \sigma^2_\xi\, \widehat{G}(k,\sigma_\phi)\,. \tag{6.47}$$

Note that the approximation in Equation 6.46 holds for $\sigma_\phi \ll 2\pi$, and the function $\widehat{G}(\omega;\sigma) := \int_{-\infty}^{\infty}\mathrm{d}\tau\, G(\tau;\sigma)\exp(-j\omega\tau) = \sqrt{2\pi}\,\sigma\exp(-\omega^2\sigma^2/2)$ is the continuous

Fourier transform of the function $G(\tau;\sigma)$.

Population-level output power spectrum

In Section 6.B.1 (Equations 6.33-6.36), I showed that, at the steady-state output, the input power S^h_{pop} is amplified by a factor A_{pop} at the first harmonic, that is,

$$S^{v^\infty}_{\text{pop}}(k) = \begin{cases} A_{\text{pop}}\, S^h_{\text{pop}}(k) & |k| = 1 \\ S^h_{\text{pop}}(k) & \text{otherwise} \end{cases} . \tag{6.48}$$

6.B.3. Single-cell power spectra

Here, I compute the single-cell power spectra of the feed-forward input h and of the steady-state output v_∞ (Section 6.4.3 and Figure 6.9).

Single-cell input power spectrum

I define the single-cell input power spectrum

$$S^h_{\text{cell}}(k) := \int_{-L/2}^{L/2} \mathrm{d}\tau\, C^h_{\text{cell}}(\tau) \exp\left(-\frac{j2\pi k\tau}{L}\right) \tag{6.49}$$

where L is the length of the track and

$$C^h_{\text{cell}}(\tau) := \Big\langle h(\phi,x)h(\phi,x+\tau)\Big\rangle_x \tag{6.50}$$

$$\overset{(6.3)}{=} \underbrace{\Big\langle g(\phi,x)g(\phi,x+\tau)\Big\rangle_x}_{=:\, C^g_{\text{cell}}(\tau)} + \underbrace{\Big\langle \xi(\phi,x)\xi(\phi,x+\tau)\Big\rangle_x}_{=:\, C^\xi_{\text{cell}}(\tau)} \quad \forall\phi\,. \tag{6.51}$$

is the single-cell autocorrelation of the feed-forward input h (Equation 6.3). In Equations 6.50 and 6.51, g is the input signal, ξ is the input noise, and the angular brackets denote statistical expectation. The function

$$C^g_{\text{cell}}(\tau) = \frac{1}{L}\int_{-L/2}^{L/2} \mathrm{d}x\, g(\phi,x)g(\phi,x+\tau) \overset{(6.3)}{=} \frac{B^2}{2}\cos(2\pi f\tau) \tag{6.52}$$

is the single-cell autocorrelation of the input signal g, and the function

$$C^\xi_{\text{cell}}(\tau) = C_\xi(0,\tau) \overset{(6.5)}{=} \sigma^2_\xi G(\tau;\sigma_x) \tag{6.53}$$

is the single-cell autocorrelation of the input noise ξ.

From Equations 6.49 and 6.51, I obtain

$$S^h_{\text{cell}}(k) = S^g_{\text{cell}}(k) + S^\xi_{\text{cell}}(k) \tag{6.54}$$

where

$$S^g_{\text{cell}}(k) := \int_{-L/2}^{L/2} \mathrm{d}\tau\, C^g_{\text{cell}}(\tau) \exp\left(-\frac{2\pi jk\tau}{L}\right) \overset{(6.52)}{=} \begin{cases} B^2 L/4 & |k| = Lf \\ 0 & \text{otherwise} \end{cases} \tag{6.55}$$

is the single-cell power spectrum of the input signal g, and

$$\begin{aligned} S^\xi_{\text{cell}}(k) &:= \int_{-L/2}^{L/2} \mathrm{d}\tau\, C^\xi_{\text{cell}}(\tau) \exp\left(-\frac{2\pi jk\tau}{L}\right) && (6.56)\\ &\overset{(6.53)}{\approx} \sigma^2_\xi \int_{-\infty}^{\infty} \mathrm{d}\tau\, G(\tau;\sigma_x) \exp\left(-\frac{2\pi jk\tau}{L}\right) && (6.57)\\ &= \sigma^2_\xi\, \widehat{G}\left(\frac{2\pi k}{L};\sigma_x\right). && (6.58) \end{aligned}$$

is the single-cell power spectrum of the input noise ξ. Note that the approximation in Equation 6.57 holds for $\sigma_x \ll L$, and the function $\widehat{G}(\omega;\sigma) := \int_{-\infty}^{\infty} \mathrm{d}\tau\, G(\tau;\sigma) \exp(-j\omega\tau) = \sqrt{2\pi}\,\sigma \exp(-\omega^2\sigma^2/2)$ is the continuous Fourier transform of $G(\tau;\sigma)$.

Steady-state output

To compute the power spectrum at the single-cell output, I first compute the steady-state output v_∞. In Section 6.B.1 (Equation 6.33), I derived the Fourier coefficients

$$\widetilde{v}^\infty_k(x) = \widetilde{h}_k(x)\, \widetilde{F}^{\text{pop}}_k \tag{6.59}$$

of the steady-state output v_∞, where

$$\widetilde{F}^{\text{pop}}_k = \begin{cases} \sqrt{A_{\text{pop}}} & |k| = 1 \\ 1 & \text{otherwise} \end{cases} \tag{6.60}$$

are the Fourier coefficients of the equivalent population feed-forward filter F^{pop} (Equations 6.35 and 6.36). By back-transforming Equation 6.59 to phase domain, I obtain

(Equation 6.31)

$$v_\infty(\phi, x) = \int_{-\pi}^{\pi} \mathrm{d}\tau\, h(\tau, x)\, F^{\mathrm{pop}}(\phi - \tau) \tag{6.61}$$

$$= \underbrace{\int_{-\pi}^{\pi} \mathrm{d}\tau\, g(\tau, x) F^{\mathrm{pop}}(\phi - \tau)}_{=:\, g^{\mathrm{out}}(\phi, x)} + \underbrace{\int_{-\pi}^{\pi} \mathrm{d}\tau\, \xi(\tau, x) F^{\mathrm{pop}}(\phi - \tau)}_{=:\, \xi^{\mathrm{out}}(\phi, x)} \tag{6.62}$$

where g^{out} is the output signal, and ξ^{out} is the output noise. Note that the output signal

$$g^{\mathrm{out}}(\phi, x) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \widetilde{g}_k(x) \widetilde{F}^{\mathrm{pop}}_k \exp(j\phi k) \overset{(6.60)}{=} \sqrt{A_{\mathrm{pop}}}\, g(\phi, x) \tag{6.63}$$

has single-cell autocorrelation

$$C^{g^{\mathrm{out}}}_{\mathrm{cell}}(\tau) = \frac{1}{L} \int_{-L/2}^{L/2} \mathrm{d}x\, g^{\mathrm{out}}(\phi, x) g^{\mathrm{out}}(\phi, x + \tau) \overset{(6.63)}{=} A_{\mathrm{pop}}\, C^{g}_{\mathrm{cell}}(\tau)\,. \tag{6.64}$$

Output noise

I now compute the autocorrelation

$$C_{\xi^{\mathrm{out}}}(\tau_\phi, \tau_x) := \left\langle \xi^{\mathrm{out}}(\phi, x)\, \xi^{\mathrm{out}}(\phi + \tau_\phi, x + \tau_x) \right\rangle_{\phi, x} \tag{6.65}$$

$$= \int_{-\pi}^{\pi} \mathrm{d}\phi\, C_\xi(\tau_\phi - \phi, \tau_x) \int_{-\pi}^{\pi} \mathrm{d}\phi'\, F^{\mathrm{pop}}(\phi') F^{\mathrm{pop}}(\phi + \phi') \tag{6.66}$$

of the output noise ξ^{out}. Equation 6.66 shows that $C_{\xi^{\mathrm{out}}}$ is obtained by convolving the input-noise autocorrelation C_ξ with the autocorrelation of the population filter F^{pop}:

$$\int_{-\pi}^{\pi} \mathrm{d}\phi'\, F^{\mathrm{pop}}(\phi') F^{\mathrm{pop}}(\phi + \phi') = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} |\widetilde{F}^{\mathrm{pop}}_k|^2 \exp(jk\phi) \tag{6.67}$$

$$\overset{(6.33)}{=} \frac{A_{\mathrm{pop}} - 1}{\pi} \cos(\phi) + \delta(\phi)\,. \tag{6.68}$$

Therefore, from Equations 6.66 and 6.68, I obtain

$$C_{\xi^{\mathrm{out}}}(\tau_\phi, \tau_x) = C_\xi(\tau_\phi, \tau_x) + \frac{A_{\mathrm{pop}} - 1}{\pi} \int_{-\pi}^{\pi} \mathrm{d}\phi\, C_\xi(\tau_\phi - \phi, \tau_x) \cos(\phi)\,. \tag{6.69}$$

I now recall the definition of the input-noise autocorrelation (Equation 6.5)

$$C_\xi(\tau_\phi, \tau_x) := \sigma_\xi^2\, G(\tau_\phi; \sigma_\phi) G(\tau_x; \sigma_x) \quad \text{with} \quad G(\tau, \sigma) := \exp\left(-\frac{\tau^2}{2\sigma^2}\right) \tag{6.70}$$

where $\sigma_\xi^2 > 0$ is the noise variance. Plugging Equation 6.70 into Equation 6.69 yields

$$\begin{aligned}
\int_{-\pi}^{\pi} \mathrm{d}\phi\, C_\xi(\tau_\phi - \phi, \tau_x)\cos(\phi) &= \sigma_\xi^2 G(\tau_x;\sigma_x)\int_{-\pi}^{\pi}\mathrm{d}\phi\, G(\tau_\phi-\phi;\sigma_\phi)\cos(\phi) && (6.71)\\
&\approx \sigma_\xi^2 G(\tau_x;\sigma_x)\int_{-\infty}^{\infty}\mathrm{d}\phi\, G(\tau_\phi-\phi;\sigma_\phi)\cos(\phi) && (6.72)\\
&= \sigma_\xi^2 G(\tau_x;\sigma_x)\cos(\tau_\phi)\sqrt{2\pi}\,\sigma_\phi \exp(-\sigma_\phi^2/2) && (6.73)\\
&= \sigma_\xi^2 G(\tau_x;\sigma_x)\cos(\tau_\phi)\widehat{G}(1;\sigma_\phi)\,. && (6.74)
\end{aligned}$$

Note that the approximation in Equation 6.72 holds for $\sigma_\phi \ll 2\pi$, and the function $\widehat{G}(\omega;\sigma) := \int_{-\infty}^{\infty}\mathrm{d}\tau\, G(\tau;\sigma)\exp(-j\omega\tau) = \sqrt{2\pi}\,\sigma\exp(-\omega^2\sigma^2/2)$ is the continuous Fourier transform of $G(\tau;\sigma)$. Finally, using Equations 6.70 and 6.74 in Equation 6.69 yields

$$C_{\xi^{\text{out}}}(\tau_\phi,\tau_x) \approx \sigma_\xi^2 G(\tau_x;\sigma_x)\left[G(\tau_\phi;\sigma_\phi) + \frac{A_{\text{pop}}-1}{\pi}\cos(\tau_\phi)\widehat{G}(1;\sigma_\phi)\right]. \tag{6.75}$$

Equation 6.75 shows that the recurrent dynamics changes the noise correlations across neurons (τ_ϕ) but not across space (τ_x).

Single-cell output power spectrum

Here, I compute the single-cell power spectrum

$$S_{\text{cell}}^{v^\infty}(k) = S_{\text{cell}}^{g^{\text{out}}}(k) + S_{\text{cell}}^{\xi^{\text{out}}}(k) \tag{6.76}$$

of the steady-state output v_∞, where $S_{\text{cell}}^{g^{\text{out}}}$ is the single-cell power spectrum of the output signal and $S_{\text{cell}}^{\xi^{\text{out}}}$ is the single-cell power spectrum of the output noise. I derive

$$S_{\text{cell}}^{g^{\text{out}}}(k) := \int_{-L/2}^{L/2}\mathrm{d}\tau\, C_{\text{cell}}^{g^{\text{out}}}(\tau)\exp\left(-\frac{j2\pi k\tau}{L}\right) \overset{(6.64)}{=} A_{\text{pop}} S_{\text{cell}}^{g}(k)\,. \tag{6.77}$$

To compute the power spectrum $S_{\text{cell}}^{\xi^{\text{out}}}$, I first compute the single-cell autocorrelation

$$\begin{aligned}
C_{\text{cell}}^{\xi^{\text{out}}}(\tau) &:= \left\langle \xi^{\text{out}}(\phi,x)\xi^{\text{out}}(\phi,x+\tau)\right\rangle_x && (6.78)\\
&\overset{(6.65)}{=} C_{\xi^{\text{out}}}(0,\tau) && (6.79)\\
&\overset{(6.75)}{\approx} \sigma_\xi^2\, G(\tau;\sigma_x)\left[1+\frac{A_{\text{pop}}-1}{\pi}\widehat{G}(1;\sigma_\phi)\right] && (6.80)\\
&\overset{(6.53)}{\approx} A_{\text{noise}}\, C_{\text{cell}}^{\xi}(\tau) && (6.81)
\end{aligned}$$

where

$$A_{\text{noise}} := 1 + \frac{A_{\text{pop}} - 1}{\pi} \widehat{G}(1; \sigma_\phi) \tag{6.82}$$

$$= 1 + (A_{\text{pop}} - 1) \sqrt{\frac{2}{\pi}} \sigma_\phi \exp\left(-\frac{\sigma_\phi^2}{2}\right). \tag{6.83}$$

Therefore, the single-cell power spectrum of the output noise reads

$$S_{\text{cell}}^{\zeta^{\text{out}}}(k) := \int_{-L/2}^{L/2} \mathrm{d}\tau \, C_{\text{cell}}^{\zeta^{\text{out}}}(\tau) \exp\left(-\frac{2\pi j k \tau}{L}\right) \overset{(6.81)}{\approx} A_{\text{noise}} \, S_{\text{cell}}^{\zeta}(k). \tag{6.84}$$

Equations 6.77 and 6.84 show that the power of the input signal is amplified by a factor A_{pop} (Equation 6.36) whereas the power of the input noise is amplified by a factor A_{noise} (Equation 6.83). Therefore, the effects of the recurrent connections on the single-cell network activity are summarized by the power spectrum of the the equivalent *single-cell feed-forward* filter F^{cell}:

$$|\widetilde{F}_k^{\text{cell}}|^2 := \frac{S_{\text{cell}}^{v^\infty}(k)}{S_{\text{cell}}^{h}(k)} = \frac{A_{\text{pop}} S_{\text{cell}}^{g}(k) + A_{\text{noise}} S_{\text{cell}}^{\zeta}(k)}{S_{\text{cell}}^{g}(k) + S_{\text{cell}}^{\zeta}(k)}. \tag{6.85}$$

Because the grid signal g has only power at the tuning frequency $k = Lf$ (Equation 6.55), Equation 6.85 can be rewritten as follows:

$$|\widetilde{F}_k^{\text{cell}}|^2 := \begin{cases} A_{\text{cell}} & k = Lf \\ A_{\text{noise}} & \text{otherwise} \end{cases} \tag{6.86}$$

where

$$A_{\text{cell}} := \frac{A_{\text{pop}} S_{\text{cell}}^{g}(Lf) + A_{\text{noise}} S_{\text{cell}}^{\zeta}(Lf)}{S_{\text{cell}}^{g}(Lf) + S_{\text{cell}}^{\zeta}(Lf)}. \tag{6.87}$$

Equation 6.86 shows that, at the single-cell level, the input power is amplified by a factor A_{cell} at the tuning harmonic $k = Lf$, and by a factor A_{noise} at all other frequencies.

Finally, from Equation 6.86, I derive

$$F^{\text{cell}}(x) = \frac{1}{L} \sum_{k=-\infty}^{\infty} \widetilde{F}_k^{\text{cell}} \exp\left(\frac{2\pi j k x}{L}\right) \tag{6.88}$$

$$= \sqrt{A_{\text{noise}}} \, \delta(x) + \frac{2}{L} (\sqrt{A_{\text{cell}}} - \sqrt{A_{\text{noise}}}) \cos(2\pi f x). \tag{6.89}$$

6.B.4. Amplification Index

The improvement of a single-cell spatial tuning is quantified by the amplification index (Equation 6.7)

$$\Psi := \frac{A_{\text{cell}}}{A_{\text{noise}}} \,. \tag{6.90}$$

A simple expression of the index Ψ is obtained by rewriting the amplification factor A_{cell} in terms of the input signal-to-noise ratio (Equations 6.55 and 6.58)

$$\text{SNR}_{\text{in}} := \frac{S^{g}_{\text{cell}}(Lf)}{S^{\xi}_{\text{cell}}(Lf)} = \frac{B^2 L/4}{\sigma^2_{\xi} \widehat{G}(2\pi f; \sigma_x)} \,. \tag{6.91}$$

By using Equation 6.91 in Equation 6.87, I derive

$$A_{\text{cell}} = \frac{A_{\text{pop}} \,\text{SNR}_{\text{in}} + A_{\text{noise}}}{\text{SNR}_{\text{in}} + 1} \,. \tag{6.92}$$

Finally, from Equations 6.90 and 6.92, I obtain

$$\Psi = \frac{A_{\text{pop}} / A_{\text{noise}} \,\text{SNR}_{\text{in}} + 1}{\text{SNR}_{\text{in}} + 1} \,. \tag{6.93}$$

Part III.

Synthesis

Chapter 7

Conclusion

Here, I summarize the main results of this thesis (Section 7.1), review experimental predictions (Section 7.2), and provide directions for future research (Section 7.3).

7.1. Thesis summary and conclusions

The goal of this thesis was to investigate the mechanisms underlying the emergence, inheritance, and amplification of grid-cell activity in the MEC. I approached this problem from a theoretical point of view. Through analytical work and simulations, I studied mathematical models of grid-cell activity that explain available empirical data and generate predictions for future experiments.

First, I summarized anatomical and physiological data on the hippocampal system, focusing on the findings that were most relevant to grid-cell activity (Chapter 2). I then critically reviewed the grid-cell literature from both an experimental (Chapter 3) and a theoretical perspective (Chapter 4). Next, I described a novel computational model for the emergence of grid-cell activity (Chapter 5). Inspired by a previous proposal by Kropff and Treves (2008), I showed that grids could emerge via a single-cell mechanism based on spatially-tuned inputs and neuronal adaptation. Compared to previous related work (e.g., Kropff and Treves, 2008; Dordek et al., 2016), my model achieved a higher level of biological realism, gave unprecedented analytical insights, and generated novel experimental predictions. Finally, I investigated how grid-cell patterns may be affected by excitatory microcircuits in the MEC (Chapter 6). I showed that grids could be inherited across neuronal populations, and that feed-forward and recurrent circuits could amplify the periodicity of grid patterns.

In conclusion, I outlined a theoretical framework explaining the origin, inheritance, and amplification of grid-cell activity. I believe that this thesis contributes to a better understanding of the cellular mechanisms underlying the neural representation of space in the medial entorhinal cortex. Future work in this direction, could possibly unveil the neuronal underpinnings of high-level cognitive functions, such as navigation

and spatial memory.

7.2. Experimental predictions

Here, I summarize the main experimental predictions of the present work (see also discussion in Chapters 5 and 6).

1. A central result of this thesis is that grids can emerge via a single-cell process driven by Hebbian synaptic plasticity and neuronal adaptation. Therefore, I predict that grids shall be impaired by interfering with synaptic plasticity during development. This prediction is in line with the fact that grid patterns are disrupted in adult animals following knock-out of NMDA receptors in pups (Dagslott et al., 2016; Gil et al., 2018). Future experiments shall characterize the plasticity of entorhinal synapses in young animals, and assess whether this plasticity depends on NMDA receptors.

2. Similarly, grids shall be impaired by interfering with neuronal adaptation during development. For testing this prediction, the biophysical mechanisms underpinning adaptation in the MEC need to be uncovered. To obtain grid scales similar to the ones observed in laboratory experiments (e.g., 30–70 cm), the model requires adaptation-recovery time constants of the order of hundreds of milliseconds. I propose that such slow adaptation currents could be mediated by HCN2–4 channels (Accili et al., 2002). Therefore, knocking out HCN2–4 channels in the MEC, could impair the emergence of grid-cell activity during development.

3. In my model, the grid spacing depended on the adaptation dynamics, the spatial scale of the input tuning, and the average exploration speed of the animal during learning (Chapter 5). Interfering with any of those variables during development shall thus affect grid periodicity during adulthood. Additionally, the model predicts that grids at larger scales shall develop faster than grids at finer scales. Because grid scale is topographically organized in the MEC (Hafting et al., 2005), I suggest that the earliest signatures of grid-cell activity shall be observed in the most ventral entorhinal areas.

4. In Chapter 6, I showed that grids can be inherited across neuronal populations via feed-forward projections, and that their spatial periodicity can be improved in this inheritance process. I thus suggest that regions where grid cells are most abundant could be functionally downstream to regions where only weaker grid tuning is found. Considering the layout of the MEC excitatory circuitry (Fuchs et al., 2016; Winterer et al., 2017), I predict that inactivating either layer II or layer III pyramidal cells shall result in a reduced grid tuning of layer II stellate cells.

5. Finally, I demonstrated that recurrent excitation can amplify grid tuning in local populations of neurons, and that a connectivity supporting this function can self-organize via synaptic plasticity during development. The functional role of recurrent connectivity in the MEC could be experimentally assessed by silencing local grid-cell populations and simultaneously recording excitatory post-synaptic currents at the feed-forward input (Lien and Scanziani, 2013). The amplification model I proposed in this thesis predicts that such inputs shall be weakly grid tuned. By contrast, a completely untuned input would speak in favor of CAN models. Additionally, the developmental maturation of recurrent connectivity (and its dependence on spatial experience) could be probed by a combining *in vivo* imaging and *in vitro* recordings in young animals (Ko et al., 2013).

7.3. Future work

Finally, I outline a few follow-up projects related to the work of this thesis.

Towards a biophysical model of grid-cell activity. In Chapter 5, I reformulated the model by Kropff and Treves (2008) in a more biologically-plausible setting. Because the Kropff-and-Treves model included an abstract network-level normalization of the neuronal activity (which induced competitive dynamics between the neurons), a first task was to show that grids could still arise via a purely single-cell process. To improve biological realism, I then modeled neural activity, synaptic plasticity, and adaptation dynamics in a spike-based setting. Yet the neuron model I chose (the Poisson neuron) remained rather simple, e.g., it abstracted from the biophysical underpinnings of spike-rate adaptation. This choice was motivated by previous work in our laboratory that has proven difficult to study a complex phenomenon such as grid-pattern formation in a detailed conductance-based neuron (Herding, 2012). I thus opted for a neuron model of intermediate complexity, which on the one hand could give insights on the spiking dynamics, and on the other hand remained analytically tractable. This allowed to derive precise mathematical requirements for grid-pattern formation in a spiking framework (Chapter 5). I believe that such requirements could now constrain more biophysically-detailed models of grid-cell activity. A possible step in this direction is to test whether grids are still obtained using an integrate-and-fire neuron model with adaptation (e.g., Liu and Wang, 2001; Brette and Gerstner, 2005).

Theta modulation and phase precession. Salient temporal features of grid-cell activity are theta-modulation and phase precession, i.e., that within grid fields spikes occur at earlier and earlier phases of the extracellular theta rhythm (Hafting et al.,

2008; Climer et al., 2013; Jeewajee et al., 2014; Newman and Hasselmo, 2014; Reifenstein et al., 2014). It is to date unclear whether phase phase precession is generated together with spatially-periodic firing—as proposed by oscillatory-interference models (e.g., O'Keefe and Burgess, 2005; Burgess, 2008; Blair et al., 2008)—or rather via a separate mechanism (e.g., Thurley et al., 2013). In the latter case, grid-cell phase precession could be inherited by afferent inputs to the MEC (Jaramillo et al., 2014), e.g., hippocampal place cells (O'Keefe and Recce, 1993) or parasubicular cells (Ebbesen et al., 2016). In this view, it would be interesting to test whether the inheritance of phase precession is compatible with the grid-cell model described in Chapter 5, and to study how phase precession could affect the formation of grid patterns (D'Albis et al., 2015; Monsalve-Mercado and Leibold, 2017).

Grids and the geometry of the environment. Experimental evidence suggests that grids are affected by the geometry of the enclosure, e.g., they align to the boundaries in quadrangular arenas (Krupic et al., 2015; Stensola et al., 2015), and they are distorted in more complex arena shapes, such as the hairpin maze (Derdikman et al., 2009) or trapezoidal environments (Krupic et al., 2015, 2018, see also Section 3.2). The grid-cell model proposed in this thesis, however, remained too abstract to account for such observations (Chapter 5). In fact, the simulations of Chapter 5 assumed periodic boundary conditions at the edges of the environment. This assumption was needed to demonstrate the basic principle of grid-pattern formation eliminating confounds due to border artefacts. However, because environmental geometry affects grid orientation and symmetry, it would be interesting to extend the grid-cell model presented here to more realistic boundary conditions.

Grids across environments. In the grid-cell model of Chapter 5, pattern formation requires feed-forward inputs that are already spatially tuned. Such a spatial tuning could be generated locally in the MEC or provided by afferent inputs, e.g., hippocampal projections to the deep entorhinal layers or parasubicular projections to the superficial layers. The underlying idea is that allocentric representations of space (e.g., place fields) could emerge prior to grid-cell tuning (Wills et al., 2010; Langston et al., 2010) by integrating egocentric sensory cues (Hartley et al., 2000; Strösslin et al., 2005; Franzius et al., 2007). Yet, because hippocampal place cells and entorhinal (non-grid) spatial cells remap across environments (Leutgeb et al., 2005; Diehl et al., 2017), grids would have to be re-learned in each novel environment anew (note that little is known about remapping in parasubiculum instead). My grid-cell model, however, requires extensive spatial exploration for learning, which contrasts to the rapid emergence of grids in novel enclosures

(Hafting et al., 2005). This problem could be solved by considering spatial inputs that are invariant to environmental context (e.g., border cells)—a possibility that needs to be tested thoroughly. Furthermore, co-modular grids undergo coherent remapping across environments, i.e., their relative spatial phases are preserved up to a common rotation (Fyhn et al., 2007; Yoon et al., 2013). Such a coherent transformation could stem from a structured recurrent connectivity akin to the one emerging in the simulations of Chapter 6 (Section 6.3). It would be thus interesting to study how feed-forward and recurrent inputs interact to generate coherent representations across environments (see also Si et al., 2012, for a similar approach).

Learning associations with self-motion cues. In Chapter 6, I showed that structured recurrent connectivity could self-organize from weakly grid-tuned inputs. I also proposed that such recurrent connections could sustain grid-like firing when the feed-forward tuning is temporarily lost, i.e., in the absence of strong or familiar sensory input. For this to happen, synaptic connections between the grid-cell network and inputs carrying self-motion signals (e.g., speed and head-direction cells) need to be learned. Self-motion inputs shall have asymmetric projections to the grid cells such that the network activity is updated coherently with the animal's movements (see Section 6.3.3, Figure 6.5 for further details). Additionally, and critically for the framework suggested in this thesis, self-motion integration shall match the spatial tuning of the feed-forward inputs, i.e., spatial (feed-forward) and self-motion (recurrent) signals shall always be aligned during exploration (contrast to Widloski and Fiete, 2014). Previous theoretical work has shown that networks integrating velocity inputs could develop in a Hebbian framework (Stringer et al., 2002b,a; Widloski and Fiete, 2014), yet it remains unclear whether an alignment between spatial and self-motion cues could also be learned in the same setting.

Chapter 8

Outlook

I want to end this thesis by addressing open questions on the functional role of grid-cell activity in animal behavior. In particular, I would like to discuss how grid cells may contribute to spatial navigation (Section 8.1), spatial memory (Section 8.2), and abstract cognition (Section 8.3).

8.1. Grid cells and spatial navigation

Since its discovery, grid-cell activity has been regarded as a neural representation of space guiding animals to navigate in their environment. In 2014, John O'Keefe and Edvard and May-Britt Moser were awarded the Nobel Prize in Physiology or Medicine for the discovery of "a comprehensive positioning system, an inner GPS, in the brain that makes it possible to orient ourselves in space." The discovery of place and grid cells was certainly a major breakthrough in systems neuroscience, but the role of these cells in spatial cognition and navigation is far from being understood. Remarkably, most of place- and grid-cell data has been collected in contexts with minimal navigational demands, such as foraging in small featureless arenas, or running on straight linear tracks. By contrast, navigation in natural environments is a highly demanding task, which requires multiple sensory cues, cognitive strategies, and brain circuits.

8.1.1. How do animals navigate?

At the behavioral level, animal navigation has fascinated ethologists since the 19th century (e.g., Wallace, 1873; Darwin, 1873; Murphy, 1873). By studying the displacement of mostly non-mammalian species (e.g., ants, honeybees, migratory birds, and sea turtles) over large spatial scales (from tens of meters to kilometers), ethologist constructed a taxonomy of navigational strategies at different levels of cognitive demand (e.g., Gallistel, 1990; Alerstam and Hedenström, 1998; Wallraff, 2005; Eichenbaum, 2017). Such strategies include *random search* (locomotion and goal recognition), *beaconing*

(orientation through observable distal cues or landmarks), *praxis* (navigation via stereotyped internally-guided movements), *route following* (guidance through familiar routes), and *survey* (navigation based on an internalized mental map of the environment). The latter, survey navigation, is the most sophisticated strategy but also the most flexible. By embedding known places and their relations in a common reference frame, the animal constructs a *cognitive map* of the environment that allows to plan novel routes in familiar spaces.

8.1.2. The cognitive map theory

The cognitive map theory was formulated by the psychologist Edward Tolman on the basis of behavioral experiments in rodents (Tolman, 1948), and it was later revived by John O'Keefe and Lynn Nadel after the discovery of place cells in the rat hippocampus (O'Keefe and Nadel, 1978). In their famous treatise, O'Keefe and Nadel (1978) reviewed an impressive number of studies relating hippocampal activity to spatial cognition, and proposed that hippocampal place cells were the neuronal instantiation of the cognitive map theorized by Tolman thirty years before. Today, not only place cells, but also grid, border, and head-direction cells are believed to support spatial navigation in mammals by creating a Tolmanian mental map of the environment.

The cognitive map theory posits the need to combine self-motion and landmark-based sensory cues to perform self-triangulation. Yet the grid-cell literature has often downplayed landmark-based navigation in favor to orientation strategies purely based on self-motion inputs—also termed *path integration* or *dead reckoning* (Moser et al., 2017). In the following, I shall briefly review current empirical knowledge on path integration, and argue that mammals such as rats or humans (where grid cells are found) are unlikely to use path integration as a primary orientation strategy.

8.1.3. Path integration

Navigation via path integration has been studied mostly in relation to *homing*, i.e. the ability of some animals to return to their nest after a tortuous outbound journey. An excellent path integrator is the desert ant, *Cataglyphis fortis*, a solitary forager that after hunting for hundreds of meters in a featureless terrain returns directly to its nest with an approximately straight trajectory. Experimental evidence suggests that during an outbound journey, the animal continuously integrates angles steered and distance traveled in a mean displacement vector that is then used for homing (Müller and Wehner, 1988). Steering angles are measured using skylight information (mainly light-polarization patterns, Wehner, 1976) whereas linear displacement is deduced from optic flow (Ronacher and Wehner, 1995) and kinesthetic signals (i.e., by counting steps, Wittlinger et al., 2006). However, even a champion of path integration such as the

desert ant switches to landmark-based navigation when familiar places are encountered (Collett et al., 1992), and follows external signals when internal and external cues are put in conflict by the experimenter (Collett et al., 1998).

These results suggests that mammals such as rodents or humans, who live in highly cue-rich environments, predominantly employ landmark-based strategies for orientation—possibly falling back to path integration only when external cues are temporarily unavailable. In fact, when tested in a laboratory setting, mammals such as hamsters, rats, dogs, and humans show surprisingly poor path-integration capabilities (Etienne et al., 1996; Loomis et al., 1993; Kim et al., 2013; Gil et al., 2018; Stangl et al., 2018; Tennant et al., 2018). Remarkably, mice cannot rely on path integration for distances longer than 2 m (Tennant et al., 2018), and blind or blindfolded humans make distance-estimation errors of the order of 10% when walking on straight lines for 2–10 m (Loomis et al., 1993).

Yet a prominent view in the field is that mammals navigate through path integration and that path integration is supported by grid-cell activity (Moser et al., 2017). A link between grid cells and path integration was suggested by early studies in which grid fields persisted during open-field foraging in the dark—indicating that grid patterns could be generated by self-motion cues only (Hafting et al., 2005). Yet grids are strongly influenced by environmental cues (e.g., Barry et al., 2007; Krupic et al., 2015, 2018; Wernle et al., 2018), and they fall apart in the dark when olfactory and somatosensory cues are removed (Chen et al., 2016; Pérez-Escobar et al., 2016, see also Section 3.2).

Gil et al. (2018) recently sought to determine the link between grid-cell activity and path integration in a more direct manner. The authors showed that ablating NMDA receptors in the retro-hippocampal region of young mice disrupts both grid-cell activity and path-integration performances in the adulthood. By contrast, head-direction, border, and place-cell selectivity remained largely unaffected by the intervention. This study demonstrated that NMDA receptors are required for the ontogenetic development of grid patterns, and that grids cells can support path integration for brief excursions.

An involvement of grid-cell activity in path integration is further supported by a recent study by Tennant et al. (2018). The authors devised an experimental paradigm to disentangle beaconing (cue-based navigation) from path-integration strategies in rodents. Mice were trained to stop at a defined location on a virtual linear track to receive rewards. In a fraction of the trials, the reward zone was clearly marked by visual cues (beaconed trials), whereas in the remaining trials no visual information was present at the reward zone (non-beaconed trials to probe path integration). Mice could learn the task in both scenarios, but in non-beaconed trials learning performances steeply dropped with distance, i.e., reward locations at 2–5 m from the starting position were correctly identified in less than $\sim$25% of the trials. By manipulating the motor-

visual gain of the virtual-reality system, the authors deduced that path integration was mainly driven by motor cues (e.g., proprioceptive feedback of motor-afferent copies) rather than optic flow. Finally, inactivation of stellate cells in MEC layer II (putative grid cells) led to substantial learning deficits in the same task, particularly for non-beaconed trials (Tennant et al., 2018).

The studies by Gil et al. (2018) and Tennant et al. (2018) indicate that grid cells contribute to self-location and spatial learning in rodents, but also suggest that the same animals cannot reliably use path integration for ethologically-relevant distances. Finally, I want to stress that even though grid cells may support path-integration strategies in mammals, there is no reason to think that grid patterns themselves are generated via self-motion inputs—as proposed by CAN models for example (McNaughton et al., 2006).

8.1.4. Possible uses of grid-cell activity for navigation

From a theoretical point of view, grid-cell activity could provide two pieces of information to the animal: its current position in the environment, and its distance relative to a reference point or goal location (e.g., Fiete et al., 2008b; Sreenivasan and Fiete, 2011; Erdem and Hasselmo, 2012; Kubie and Fenton, 2012; Mathis et al., 2012a,b; Stemmler et al., 2015; Wei et al., 2015; Bush et al., 2015); see also (Herz et al., 2017) for a recent review.

Stemmler et al. (2015) suggested a biologically-plausible mechanism to infer animal location from the activity of a population of grid cells. Their method is based on the computation of population vectors—an approach that was initially used to decode movement direction in motor cortex (Georgopoulos et al., 1986); see (Pouget et al., 2000) for a review. The model assumes that a downstream neuron (the decoder) receives convergent input from a population of grid cells with different spatial scales and phase offsets. With these inputs, a rough estimate of the animal location is first obtained by weighting the activity of grid cells at the largest scale with their corresponding spatial phase. Such estimate, however, is inherently ambiguous owing to the periodicity of the input tuning. To solve this problem, population vectors are computed from grids at increasingly smaller scales, which are then combined to obtain a precise position estimate at the output (Stemmler et al., 2015). Because population-vector coding requires several grids at each scale, the model inherently predicts the existence of discrete functional modules—which are indeed found experimentally (Stensola et al., 2012). Additionally, the model predicts an optimal scale ratio of 1.5, which is in agreement with empirical data (Stensola et al., 2012; Krupic et al., 2015); see also (Wei et al., 2015) for a similar prediction using different methods.

This work demonstrates that grid-cell activity could be used for self-location. But

why should the brain represent space—a linear quantity—in periodic coordinates? Theoretical studies show that periodic multi-scale representations of space vastly outperform their unimodal counterparts, i.e., the range of distances that can be coded by a population of grid cells increases exponentially with the number of neurons N, whereas distances coded by place cells scale only linearly with N (Sreenivasan and Fiete, 2011); see also (Mathis et al., 2012a,b) for complementary calculations on spatial resolution. Remarkably, a handful of grid modules with spacings ranging from tens of centimeters to a few meters (Stensola et al., 2012) could theoretically code distances up to several kilometers (Fiete et al., 2008a; Stemmler et al., 2015)—distances that are behaviorally relevant for wild rats (Taylor, 1978). By contrast, a similar spatial coverage cannot be obtained with place-field sizes typically measured in laboratory experiments. Yet the spatial range of a place code could be extended with wider firing fields or multiple firing locations per cell (Geva-Sagiv et al., 2015), both of which have been indeed observed in large environments (Kjelstrup et al., 2008; Rich et al., 2014; Eliav et al., 2017).

Finally, it was proposed that grid cells could guide animals towards goal locations via vector-based strategies (Bush et al., 2015; Stemmler et al., 2015). Interestingly, vectorial representations of goals have been experimentally found in the bat (Sarel et al., 2017), where grid cells are also found (Yartsev et al., 2011). Vector-based navigation requires computing translation vectors between two arbitrary locations—a problem which is inverse to that of path integration, i.e., the estimation of current location based on displacement vectors. Bush et al. (2015) proposed an algorithm to compute translation vectors from grid-cell activity and suggested multiple network architectures that could implement this algorithm. Along the same lines, Stemmler et al. (2015) devised a method to compute goal locations within the same population-coding scheme they propose for self-location. Both approaches, however, remain at a rather abstract level of description, and it is to date unclear whether such algorithms are actually implemented in the brain.

In summary, both experimental and theoretical studies suggest a role of grid cells in spatial orientation. A major challenge for future research will be to to determine whether (and how) mammals use grid-cell activity to solve complex navigational tasks in their natural environment.

8.2. Grid cells and spatial memory

Since the case of the amnesic patient H.M., more than 70 years of clinical, behavioral, and physiological research pointed at the hippocampus and the entorhinal cortex as critical brain regions for the encoding and consolidation of episodic memories (Squire, 1992; Nadel and Moscovitch, 1997; Eichenbaum, 2000; Roy and Tonegawa,

2017; Ólafsdóttir et al., 2018). However, it is still unknown how place and grid cells contribute to memory processes.

In the last years, optogenetic manipulations in rodents have demonstrated a causal link between hippocampal place-cell activity and the formation of memories with a spatial context (see e.g., Roy and Tonegawa, 2017, for a review). Furthermore, the phenomenon of hippocampal 'replay' (i.e., the offline reactivation of assembly sequences related to spatial experiences) suggests an involvement of place cells in systems memory consolidation, i.e., the transfer of memories from the hippocampus to the neocortex for a more permanent storage (see e.g., Ólafsdóttir et al., 2018, for a review).

Much less is known, however, about the role of grid-cell activity in spatial memory. Two recent studies have shown that grid cells—like place cells—replay previously experienced trajectories during rest (Ólafsdóttir et al., 2016; O'Neill et al., 2017). Ólafsdóttir et al. (2016) found a coordinated replay between hippocampal place cells in CA1 and grid cells in the deep entorhinal layers, whereas O'Neill et al. (2017) found that grid cells in the superficial layers replay independently from the hippocampus. Although these studies suggest an involvement of grid-cell activity in spatial memory, a causal link between the two has not been found yet.

8.3. Grid cells beyond physical space

Grid cells have been largely studied in relation to animal location in physical space, but periodic representations are found in non-spatial contexts too. For example in rodents, grid cells represent distances in both space and in time (Kraus et al., 2015), and cells with periodic firing fields during locomotion often display multiple receptive fields in acoustic space (Aronov et al., 2017). In primates, entorhinal neurons show periodic tuning as a function of gaze location (Killian et al., 2012; Meister and Buffalo, 2018; Julian et al., 2018; Nau et al., 2018) and locus of attention (Wilming et al., 2018). In humans, grid signals are observed during both virtual navigation (Doeller et al., 2010; Jacobs et al., 2013) and imagined navigation of the same paths (Bellmund et al., 2016; Horner et al., 2016), suggesting a role of grid-cell activity in trajectory planning and decision making. Finally, grid codes have been found to underlie abstract navigation in two-dimensional conceptual spaces (Constantinescu et al., 2016).

Altogether these studies indicate that grid patterns may represent behvioral variables that go far beyond the spatial domain. Therefore, the study of grid-cell activity has provided us with a unique opportunity to understand high-level computation in the cortex. It will be exciting to see whether this line of research could give us further mechanistic insights into complex brain functions such as navigation, memory, decision making, and abstract cognition.

Bibliography

Abbott, L. and Blum, K. (1996) Functional significance of long-term potentiation for sequence learning and prediction. *Cereb. Cortex*, 6:406–416. CrossRef

Accili, E. A., Proenza, C., Baruscotti, M., and DiFrancesco, D. (2002) From funny current to HCN channels: 20 years of excitation. *Physiology*, 17:32–37. CrossRef

Agster, K. L. and Burwell, R. D. (2009) Cortical efferents of the perirhinal, postrhinal, and entorhinal cortices of the rat. *Hippocampus*, 19:1159–1186. CrossRef

Alerstam, T. and Hedenström, A. (1998) The development of bird migration theory. *J. Avian Biol.*, 29:343–369. CrossRef

Alonso, A., De Curtis, M., and Llinas, R. (1990) Postsynaptic Hebbian and non-Hebbian long-term potentiation of synaptic efficacy in the entorhinal cortex in slices and in the isolated adult guinea pig brain. *Proc. Nat. Acad. Sci. U.S.A.*, 87:9280–9284. CrossRef

Alonso, A. and Klink, R. (1993) Differential electroresponsiveness of stellate and pyramidal-like cells of medial entorhinal cortex layer II. *J. Neurophysiol.*, 70:128–143. CrossRef

Alonso, A. and Llinás, R. R. (1989) Subthreshold Na^+-dependent theta-like rhythmicity in stellate cells of entorhinal cortex layer ii. *Nature*, 342:175–177. CrossRef

Amaral, D. and Witter, M. (1989) The three-dimensional organization of the hippocampal formation: a review of anatomical data. *Neuroscience*, 31:571–591. CrossRef

Amari, S.-i. (1977) Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.*, 27:77–87. CrossRef

Andersen, P., Bliss, T., and Skrede, K. K. (1971) Lamellar organization of hippocampal excitatory pathways. *Exp. Brain Res.*, 13:222–238. CrossRef

Annese, J., Schenker-Ahmed, N. M., Bartsch, H., Maechler, P., Sheh, C., Thomas, N., Kayano, J., Ghatan, A., Bresler, N., Frosch, M., Klaming, R., and Corkin, S. (2014) Postmortem examination of patient HM's brain based on histological sectioning and digital 3d reconstruction. *Nat. Commun.*, 5:3122. CrossRef

Armstrong, C., Wang, J., Yeun Lee, S., Broderick, J., Bezaire, M. J., Lee, S.-H., and Soltesz, I. (2016) Target-selectivity of parvalbumin-positive interneurons in layer ii of medial entorhinal cortex in normal and epileptic animals. *Hippocampus*, 26:779–793. CrossRef

Aronov, D., Nevers, R., and Tank, D. W. (2017) Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit. *Nature*, 543:719–722. CrossRef

Azizi, A. H., Schieferstein, N., and Cheng, S. (2014) The transformation from grid cells to place cells is robust to noise in the grid pattern. *Hippocampus*, 24:912–919. CrossRef

Barnes, C. A., McNaughton, B. L., Mizumori, S. J., Leonard, B. W., and Lin, L.-H. (1990) Chapter comparison of spatial and temporal characteristics of neuronal activity in sequential stages of hippocampal processing. *Prog. Brain Res.*, 83:287–300. CrossRef

Barry, C., Ginzberg, L. L., O'Keefe, J., and Burgess, N. (2012a) Grid cell firing patterns signal environmental novelty by expansion. *Proc. Nat. Acad. Sci. U.S.A.*, 109:17687–17692. CrossRef

Barry, C., Hayman, R., Burgess, N., and Jeffery, K. J. (2007) Experience-dependent rescaling of entorhinal grids. *Nat. Neurosci.*, 10:682–684. CrossRef

Barry, C., Heys, J. G., and Hasselmo, M. E. (2012b) Possible role of acetylcholine in regulating spatial novelty effects on theta rhythm and grid cells. *Front. Neural Circuits*, 6:5. CrossRef

Barry, C., Lever, C., Hayman, R., Hartley, T., Burton, S., O'Keefe, J., Jeffery, K., and Burgess, N. (2006) The boundary vector cell model of place cell firing and spatial memory. *Rev. Neurosci.*, 17:71–97. CrossRef

Beckmann, P. (1962) Statistical distribution of the amplitude and phase of a multiply scattered field. *J. Res. Natl. Bur. Stand.*, 66D:231–240. CrossRef

Beed, P., Bendels, M. H., Wiegand, H. F., Leibold, C., Johenning, F. W., and Schmitz, D. (2010) Analysis of excitatory microcircuitry in the medial entorhinal cortex reveals cell-type-specific differences. *Neuron*, 68:1059–1066. CrossRef

Beed, P., Gundlfinger, A., Schneiderbauer, S., Song, J., Böhm, C., Burgalossi, A., Brecht, M., Vida, I., and Schmitz, D. (2013) Inhibitory gradient along the dorsoventral axis in the medial entorhinal cortex. *Neuron*, 79:1197–1207. CrossRef

Bellmund, J. L., Deuker, L., Schröder, T. N., and Doeller, C. F. (2016) Grid-cell representations in mental simulation. *eLife*, 5:e17089. CrossRef

Ben-Yishai, R., Bar-Or, R. L., and Sompolinsky, H. (1995) Theory of orientation tuning in visual cortex. *Proc. Nat. Acad. Sci. U.S.A.*, 92:3844–3848. CrossRef

Bienenstock, E. L., Cooper, L. N., and Munro, P. W. (1982) Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J. Neurosci.*, 2:32–48. CrossRef

Blair, H. T. (1996) Simulation of a thalamocortical circuit for computing directional heading in the rat. In *Adv. Neural Inf. Process. Syst.*, pages 152–158.

Blair, H. T., Gupta, K., and Zhang, K. (2008) Conversion of a phase-to a rate-coded position signal by a three-stage model of theta cells, grid cells, and place cells. *Hippocampus*, 18:1239–1255. CrossRef

Boccara, C. N., Sargolini, F., Thoresen, V. H., Solstad, T., Witter, M. P., Moser, E. I., and Moser, M.-B. (2010) Grid cells in pre-and parasubiculum. *Nat. Neurosci.*, 13:987–994. CrossRef

Bonnevie, T., Dunn, B., Fyhn, M., Hafting, T., Derdikman, D., Kubie, J. L., Roudi, Y., Moser, E. I., and Moser, M.-B. (2013) Grid cells require excitatory drive from the hippocampus. *Nat. Neurosci.*, 16:309–317. CrossRef

Borckmans, P., Dewel, G., De Wit, A., Dulos, E., Boissonade, J., Gauffre, F., and De Kepper, P. (2002) Diffusive instabilities and chemical reactions. *Int. J. Bifurc. Chaos*, 12:2307–2332. CrossRef

Bostock, E., Muller, R. U., and Kubie, J. L. (1991) Experience-dependent modifications of hippocampal place cell firing. *Hippocampus*, 1:193–205. CrossRef

Brandon, M. P., Bogaard, A. R., Libby, C. P., Connerney, M. A., Gupta, K., and Hasselmo, M. E. (2011) Reduction of theta rhythm dissociates grid cell spatial periodicity from directional tuning. *Science*, 332:595–599. CrossRef

Brette, R. and Gerstner, W. (2005) Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J. Neurophysiol.*, 94:3637–3642. CrossRef

Brun, V. H., Leutgeb, S., Wu, H.-Q., Schwarcz, R., Witter, M. P., Moser, E. I., and Moser, M.-B. (2008) Impaired spatial representation in CA1 after lesion of direct input from entorhinal cortex. *Neuron*, 57:290–302. CrossRef

Brun, V. H., Otnæss, M. K., Molden, S., Steffenach, H.-A., Witter, M. P., Moser, M.-B., and Moser, E. I. (2002) Place cells and place recognition maintained by direct entorhinal-hippocampal circuitry. *Science*, 296:2243–2246. CrossRef

Buetfering, C., Allen, K., and Monyer, H. (2014) Parvalbumin interneurons provide grid cell-driven recurrent inhibition in the medial entorhinal cortex. *Nat. Neurosci.*, 17:710–718. CrossRef

Burak, Y. and Fiete, I. (2006) Do we understand the emergent dynamics of grid cell activity? *J. Neurosci.*, 26:9352–9354. CrossRef

Burak, Y. and Fiete, I. R. (2009) Accurate path integration in continuous attractor network models of grid cells. *PLOS Comput. Biol.*, 5:1–16. CrossRef

Burgess, N. (2008) Grid cells and theta as oscillatory interference: theory and predictions. *Hippocampus*, 18:1157–1174. CrossRef

Burgess, N., Barry, C., and O'Keefe, J. (2007) An oscillatory interference model of grid cell firing. *Hippocampus*, 17:801–812. CrossRef

Burgess, N., Maguire, E. A., and O'Keefe, J. (2002) The human hippocampus and spatial and episodic memory. *Neuron*, 35:625–641. CrossRef

Burwell, R. D. (2000) The parahippocampal region: corticocortical connectivity. *Annu. N. Y. Acad. Sci.*, 911:25–42. CrossRef

Burwell, R. D. and Agster, K. L. (2008) Anatomy of the hippocampus and the declarative memory system. In *Learning and Memory: A Comprehensive Reference*, volume 3, pages 47 – 66. Elsevier. CrossRef

Bush, D., Barry, C., and Burgess, N. (2014) What do grid cells contribute to place cell firing? *Trends Neurosci.*, 37:136–145. CrossRef

Bush, D., Barry, C., Manson, D., and Burgess, N. (2015) Using grid cells for navigation. *Neuron*, 87:507–520. CrossRef

Bush, D. and Burgess, N. (2014) A hybrid oscillatory interference/continuous attractor network model of grid cell firing. *J. Neurosci.*, 34:5065–5079. CrossRef

Buzsáki, G. (2002) Theta oscillations in the hippocampus. *Neuron*, 33:325–340. CrossRef

Buzsáki, G. and Moser, E. I. (2013) Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nat. Neurosci.*, 16:130–138. CrossRef

Caballero-Bleda, M. and Witter, M. P. (1993) Regional and laminar organization of projections from the presubiculum and parasubiculum to the entorhinal cortex: an anterograde tracing study in the rat. *J. Comp. Neurol.*, 328:115–129. CrossRef

Caballero-Bleda, M. and Witter, M. P. (1994) Projections from the presubiculum and the parasubiculum to morphologically characterized entorhinal-hippocampal projection neurons in the rat. *Exp. Brain Res.*, 101:93–108. CrossRef

Cacucci, F., Lever, C., Wills, T. J., Burgess, N., and O'Keefe, J. (2004) Theta-modulated place-by-direction cells in the hippocampal formation in the rat. *J. Neurosci.*, 24:8265–8277. CrossRef

Cajal, S. R. Y. (1893) Estructura del asta de ammon y fascia dentata. In *Anales de la Sociedad Española de Historia Natural*, volume 22, pages 53–114. Est. tip. de Fortanet.

Camperi, M. and Wang, X.-J. (1998) A model of visuospatial working memory in prefrontal cortex: recurrent network and cellular bistability. *J. Comput. Neurosci.*, 5:383–405.

Cannon, S. C., Robinson, D. A., and Shamma, S. (1983) A proposed neural network for the integrator of the oculomotor system. *Biol. Cybern.*, 49:127–136. CrossRef

Canto, C. B., Koganezawa, N., Beed, P., Moser, E. I., and Witter, M. P. (2012) All layers of medial entorhinal cortex receive presubicular and parasubicular inputs. *J. Neurosci.*, 32:17620–17631. CrossRef

Canto, C. B., Wouterlood, F. G., and Witter, M. P. (2008) What does the anatomical organization of the entorhinal cortex tell us? *Neural Plast.*, 2008:381243. CrossRef

Carandini, M. and Ringach, D. L. (1997) Predictions of a recurrent model of orientation selectivity. *Vision Res.*, 37:3061–3071. CrossRef

Carpenter, F., Manson, D., Jeffery, K., Burgess, N., and Barry, C. (2015) Grid cells form a global representation of connected environments. *Curr. Biol.*, 25:1176–1182. CrossRef

Castro, L. and Aguiar, P. (2014) A feedforward model for the formation of a grid field where spatial information is provided solely from place cells. *Biol. Cybern.*, 108:133–143. CrossRef

Chen, G., Manson, D., Cacucci, F., and Wills, T. J. (2016) Absence of visual input results in the disruption of grid cell firing in the mouse. *Curr. Biol.*, 26:2335–2342. CrossRef

Cheng, S. and Frank, L. M. (2011) The structure of networks that produce the transformation from grid cells to place cells. *Neuroscience*, 197:293–306. CrossRef

Climer, J. R., Newman, E. L., and Hasselmo, M. E. (2013) Phase coding by grid cells in unconstrained environments: two-dimensional phase precession. *Eur. J. Neurosci.*, 38:2526–2541. CrossRef

Codling, E. A., Plank, M. J., and Benhamou, S. (2008) Random walk models in biology. *J. R. Soc. Interface*, 5:813–834. CrossRef

Colgin, L. L. (2016) Rhythms of the hippocampal network. *Nat. Rev. Neurosci.*, 17:239–249. CrossRef

Collett, M., Collett, T. S., Bisch, S., and Wehner, R. (1998) Local and global vectors in desert ant navigation. *Nature*, 394:269–272. CrossRef

Collett, T. S., Dillmann, E., Giger, A., and Wehner, R. (1992) Visual landmarks and route following in desert ants. *J. Comp. Physiol. A*, 170:435–442. CrossRef

Conklin, J. and Eliasmith, C. (2005) A controlled attractor network model of path integration in the rat. *J. Comput. Neurosci.*, 18:183–203. CrossRef

Connors, B., Gutnick, M., and Prince, D. (1982) Electrophysiological properties of neocortical neurons in vitro. *J. Neurophysiol.*, 48:1302–1320. CrossRef

Constantinescu, A. O., O'Reilly, J. X., and Behrens, T. E. (2016) Organizing conceptual knowledge in humans with a gridlike code. *Science*, 352:1464–1468. CrossRef

Corkin, S. (2002) What's new with the amnesic patient hm? *Nat. Rev. Neurosci.*, 3:153–160.

Corkin, S., Amaral, D. G., González, R. G., Johnson, K. A., and Hyman, B. T. (1997) HM's medial temporal lobe lesion: findings from magnetic resonance imaging. *J. Neurosci.*, 17:3964–3979.

Couey, J. J., Witoelar, A., Zhang, S.-J., Zheng, K., Ye, J., Dunn, B., Czajkowski, R., Moser, M.-B., Moser, E. I., Roudi, Y., et al. (2013) Recurrent inhibitory circuitry as a mechanism for grid formation. *Nat. Neurosci.*, 16:318–324. CrossRef

Dagslott, N., Donato, F., Høydal, Ø. A., Waaga, T., Moser, M.-B., and Moser, E. I. (2016) Disrupted spatial representation following knock-out of NMDA receptors in the medial entorhinal cortex. In *Soc. Neurosci. Abstr. (183.08)*.

D'Albis, T., Jaramillo, J., and Kempter, R. (2015) Inheritance of place fields in the hippocampus through Hebbian learning. *Neural Comput.*, 27:1624–1672. CrossRef

D'Albis, T. and Kempter, R. (2017) A single-cell spiking model for the origin of grid-cell patterns. *PLOS Comput. Biol.*, 13:e1005782. CrossRef

Darwin, C. (1873) Perception in the lower animals. *Nature*, 7:360. CrossRef

Davachi, L. and Dobbins, I. G. (2008) Declarative memory. *Curr. Dir. Psychol. Sci.*, 17:112–118. CrossRef

Davis, G. W. and Goodman, C. S. (1998) Synapse-specific control of synaptic efficacy at the terminals of a single neuron. *Nature*, 392:82–86. CrossRef

Dayan, P. and Abbott, L. F. (2001) *Theoretical neuroscience*, volume 806 Cambridge, MA: MIT Press.

de Curtis, M. and Llinas, R. R. (1993) Entorhinal cortex long-term potentiation evoked by theta-patterned stimulation of associative fibers in the isolated in vitro guinea pig brain. *Brain Res.*, 600:327–330. CrossRef

Deng, P.-Y. and Lei, S. (2007) Long-term depression in identified stellate neurons of juvenile rat entorhinal cortex. *J. Neurphysiol.*, 97:727–737. CrossRef

Derdikman, D., Whitlock, J. R., Tsao, A., Fyhn, M., Hafting, T., Moser, M.-B., and Moser, E. I. (2009) Fragmentation of grid cell maps in a multicompartment environment. *Nat. Neurosci.*, 12:1325–1332. CrossRef

Deshmukh, S. S. and Knierim, J. J. (2011) Representation of non-spatial and spatial information in the lateral entorhinal cortex. *Front. Behav. Neurosci.*, 5:69. CrossRef

Dhillon, A. and Jones, R. S. (2000) Laminar differences in recurrent excitatory transmission in the rat entorhinal cortex in vitro. *Neurosci.*, 99:413–422. CrossRef

Dickson, C. T., Magistretti, J., Shalinsky, M. H., Fransén, E., Hasselmo, M. E., and Alonso, A. (2000) Properties and role of I_h in the pacing of subthreshold oscillations in entorhinal cortex layer II neurons. *J. Neurphysiol.*, 83:2562–2579. CrossRef

Diehl, G. W., Hon, O. J., Leutgeb, S., and Leutgeb, J. K. (2017) Grid and nongrid cells in medial entorhinal cortex represent spatial location and environmental features with complementary coding schemes. *Neuron*, 94:83–92. CrossRef

Doeller, C. F., Barry, C., and Burgess, N. (2010) Evidence for grid cells in a human memory network. *Nature*, 463:657–661. CrossRef

Domnisoru, C., Kinkhabwala, A. A., and Tank, D. W. (2013) Membrane potential dynamics of grid cells. *Nature*, 495:199–204. CrossRef

Donato, F., Jacobsen, R. I., Moser, M.-B., and Moser, E. I. (2017) Stellate cells drive maturation of the entorhinal-hippocampal circuit. *Science*, 3:eaai8178. CrossRef

Dordek, Y., Soudry, D., Meir, R., and Derdikman, D. (2016) Extracting grid cell characteristics from place cell inputs using non-negative principal component analysis. *eLife*, 5:e10094. CrossRef

Douglas, R. J., Koch, C., Mahowald, M., Martin, K., and Suarez, H. H. (1995) Recurrent excitation in neocortical circuits. *Science*, 269:981–985. CrossRef

Dunn, B., Mørreaunet, M., and Roudi, Y. (2015) Correlations and functional connections in a population of grid cells. *PLOS Comput. Biol.*, 11:e1004052. CrossRef

Ebbesen, C. L., Reifenstein, E. T., Tang, Q., Burgalossi, A., Ray, S., Schreiber, S., Kempter, R., and Brecht, M. (2016) Cell type-specific differences in spike timing and spike shape in the rat parasubiculum and superficial medial entorhinal cortex. *Cell Rep.*, 16:1005–1015. CrossRef

Eichenbaum, H. (1993) *Memory, amnesia, and the hippocampal system* MIT press. CrossRef

Eichenbaum, H. (2000) A cortical–hippocampal system for declarative memory. *Nat. Rev. Neurosci.*, 1:41–50. CrossRef

Eichenbaum, H. (2017) The role of the hippocampus in navigation is memory. *J. Neurophysiol.*, 117:1785–1796. CrossRef

Eichenbaum, H., Kuperstein, M., Fagan, A., and Nagode, J. (1987) Cue-sampling and goal-approach correlates of hippocampal unit activity in rats performing an odor-discrimination task. *J. Neurosci.*, 7:716–732.

Ekstrom, A. D. and Ranganath, C. (2017) Space, time and episodic memory: the hippocampus is all over the cognitive map. *Hippocampus*, 00:1–8. CrossRef

Eliav, T., Las, L., and Ulanovsky, N. (2017) Representation of large-scale spaces in the hippocampus of flying bats. In *Soc. Neurosci. Abstr. (523.06)*.

Epstein, R. A., Patai, E. Z., Julian, J. B., and Spiers, H. J. (2017) The cognitive map in humans: spatial navigation and beyond. *Nat. Neurosci.*, 20:1504–1513. CrossRef

Erchova, I., Kreck, G., Heinemann, U., and Herz, A. (2004) Dynamics of rat entorhinal cortex layer ii and iii cells: characteristics of membrane potential resonance at rest predict oscillation properties near threshold. *J. Physiol.*, 560:89–110. CrossRef

Erdem, U. M. and Hasselmo, M. (2012) A goal-directed spatial navigation model using forward trajectory planning based on grid cells. *Eur. J. Neurosci.*, 35:916–931. CrossRef

Ermentrout, G. B. and Cowan, J. D. (1979) A mathematical theory of visual hallucination patterns. *Biol. Cybern.*, 34:137–150. CrossRef

Etienne, A. S. and Jeffery, K. J. (2004) Path integration in mammals. *Hippocampus*, 14:180–192. CrossRef

Etienne, A. S., Maurer, R., and Séguinot, V. (1996) Path integration in mammals and its interaction with visual landmarks. *J. Exp. Biolog.*, 199:201–209.

Faugeras, O., Grimbert, F., and Slotine, J.-J. (2008) Absolute stability and complete synchronization in a class of neural fields models. *SIAM J. Appl. Math.*, 69:205–250. CrossRef

Fenton, A. A., Kao, H.-Y., Neymotin, S. A., Olypher, A., Vayntrub, Y., Lytton, W. W., and Ludvig, N. (2008) Unmasking the ca1 ensemble place code by exposures to small and large environments: more place cells and multiple, irregularly arranged, and expanded place fields in the larger space. *J. Neurosci.*, 28:11250–11262. CrossRef

Fiete, I. R., Burak, Y., and Brookings, T. (2008a) What grid cells convey about rat location. *J. Neurosci.*, 28:6858–6871. CrossRef

Fiete, I. R., Burak, Y., and Brookings, T. (2008b) What grid cells convey about rat location. *J. Neurosci.*, 28:6858–6871. CrossRef

Frank, L. M., Brown, E. N., and Wilson, M. (2000) Trajectory encoding in the hippocampus and entorhinal cortex. *Neuron*, 27:169–178. CrossRef

Franzius, M., Sprekeler, H., and Wiskott, L. (2007) Slowness and sparseness lead to place, head-direction, and spatial-view cells. *PLOS Comput. Biol.*, 3:e166. CrossRef

Fried, I., MacDonald, K. A., and Wilson, C. L. (1997) Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron*, 18:753–765. CrossRef

Fuchs, E. C., Neitz, A., Pinna, R., Melzer, S., Caputi, A., and Monyer, H. (2016) Local and distant input controlling excitation in layer II of the medial entorhinal cortex. *Neuron*, 89:194–208. CrossRef

Fuhs, M. C. and Touretzky, D. S. (2006) A spin glass model of path integration in rat medial entorhinal cortex. *J. Neurosci.*, 26:4266–4276. CrossRef

Fujimaru, Y. and Kosaka, T. (1996) The distribution of two calcium binding proteins, calbindin D-28K and parvalbumin, in the entorhinal cortex of the adult mouse. *Neurosci. Res.*, 24:329–343. CrossRef

Fyhn, M., Hafting, T., Treves, A., Moser, M.-B., and Moser, E. I. (2007) Hippocampal remapping and grid realignment in entorhinal cortex. *Nature*, 446:190–194. CrossRef

Fyhn, M., Molden, S., Witter, M. P., Moser, E. I., and Moser, M.-B. (2004) Spatial representation in the entorhinal cortex. *Science*, 305:1258–1264. CrossRef

Gallistel, C. R. (1990) *The organization of learning*, volume 336 MIT press Cambridge, MA.

Galtier, M. N., Faugeras, O. D., and Bressloff, P. C. (2012) Hebbian learning of recurrent connections: a geometrical perspective. *Neural Comput.*, 24:2346–2383. CrossRef

Garden, D. L., Dodson, P. D., O'Donnell, C., White, M. D., and Nolan, M. F. (2008) Tuning of synaptic integration in the medial entorhinal cortex to the organization of grid cell firing fields. *Neuron*, 60:875–889. CrossRef

Georgopoulos, A. P., Schwartz, A. B., and Kettner, R. E. (1986) Neuronal population coding of movement direction. *Science*, 233:1416–1419. CrossRef

Georgopoulos, A. P., Taira, M., and Lukashin, A. (1993) Cognitive neurophysiology of the motor cortex. *Science*, 260:47–47. CrossRef

Gerstner, W., Kempter, R., Van Hemmen, J., and Wagner, H. (1996) A neuronal learning rule for sub-millisecond temporal coding. *Nature*, 383:76–78. CrossRef

Geva-Sagiv, M., Las, L., Yovel, Y., and Ulanovsky, N. (2015) Spatial cognition in bats and rats: from sensory acquisition to multiscale maps and navigation. *Nat. Rev. Neurosci.*, 16:94–108. CrossRef

Gil, M., Ancau, M., Schlesiger, M. I., Neitz, A., Allen, K., De Marco, R. J., and Monyer, H. (2018) Impaired path integration in mice with disrupted grid cell firing. *Nat. Neurosci.*, 21:81–91. CrossRef

Giocomo, L. M. and Hasselmo, M. E. (2008a) Computation by oscillations: implications of experimental data for theoretical models of grid cells. *Hippocampus*, 18:1186–1199. CrossRef

Giocomo, L. M. and Hasselmo, M. E. (2008b) Time constants of h current in layer II stellate cells differ along the dorsal to ventral axis of medial entorhinal cortex. *J. Neurosci.*, 28:9414–9425. CrossRef

Giocomo, L. M., Hussaini, S. A., Zheng, F., Kandel, E. R., Moser, M.-B., and Moser, E. I. (2011) Grid cells use HCN1 channels for spatial scaling. *Cell*, 147:1159–1170. CrossRef

Giocomo, L. M., Zilli, E. A., Fransén, E., and Hasselmo, M. E. (2007) Temporal frequency of subthreshold oscillations scales with entorhinal grid cell field spacing. *Science*, 315:1719–1722. CrossRef

Goldman, M., Compte, A., and Wang, X. (2010) Neural integrator models. In *Encyclopedia of neuroscience*. Elsevier Ltd. CrossRef

Gonzalez-Sulser, A., Parthier, D., Candela, A., McClure, C., Pastoll, H., Garden, D., Sürmeli, G., and Nolan, M. F. (2014) GABAergic projections from the medial septum selectively inhibit interneurons in the medial entorhinal cortex. *J. Neurosci.*, 34:16739–16743. CrossRef

Guanella, A., Kiper, D., and Verschure, P. (2007) A model of grid cells based on a twisted torus topology. *Int. J. Neural Syst.*, 17:231–240. CrossRef

Haas, J. S., Dorval Ii, A. D., and White, J. A. (2007) Contributions of I_h to feature selectivity in layer II stellate cells of the entorhinal cortex. *J. Comput. Neurosci.*, 22:161–171. CrossRef

Haas, J. S. and White, J. A. (2002) Frequency selectivity of layer II stellate cells in the medial entorhinal cortex. *J. Neurophysiol.*, 88:2422–2429. CrossRef

Hafting, T., Fyhn, M., Bonnevie, T., Moser, M.-B., and Moser, E. I. (2008) Hippocampus-independent phase precession in entorhinal grid cells. *Nature*, 453:1248–1252. CrossRef

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., and Moser, E. I. (2005) Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436:801–806. CrossRef

Hardcastle, K., Ganguli, S., and Giocomo, L. M. (2015) Environmental boundaries as an error correction mechanism for grid cells. *Neuron*, 86:827–839. CrossRef

Hargreaves, E. L., Rao, G., Lee, I., and Knierim, J. J. (2005) Major dissociation between medial and lateral entorhinal input to dorsal hippocampus. *Science*, 308:1792–1794. CrossRef

Hargreaves, E. L., Yoganarasimha, D., and Knierim, J. J. (2007) Cohesiveness of spatial and directional representations recorded from neural ensembles in the anterior thalamus, parasubiculum, medial entorhinal cortex, and hippocampus. *Hippocampus*, 17:826–841. CrossRef

Hartley, T., Burgess, N., Lever, C., Cacucci, F., and O'Keefe, J. (2000) Modeling place fields in terms of the cortical inputs to the hippocampus. *Hippocampus*, 10:369–379. CrossRef

Hartley, T., Lever, C., Burgess, N., and O'Keefe, J. (2014) Space in the brain: how the hippocampal formation supports spatial cognition. *Phil. Trans. R. Soc. B*, 369:20120510. CrossRef

Hasselmo, M. E. (2011) *How we remember: brain mechanisms of episodic memory* MIT press. CrossRef

Hasselmo, M. E. (2014) Neuronal rebound spiking, resonance frequency and theta cycle skipping may contribute to grid cell firing in medial entorhinal cortex. *Phil. Trans. R. Soc. B*, 369:20120523. CrossRef

Hayman, R., Chakraborty, S., Anderson, M. I., and Jeffery, K. J. (2003) Context-specific acquisition of location discrimination by hippocampal place cells. *Eur. J. Neurosci.*, 18:2825–2834. CrossRef

Herding, J. (2012) Development of grid cells : Towards a biologically plausible single-neuron model in a hebbian learning framework. Master's thesis, Humboldt Universität zu Berlin.

Herz, A. V., Mathis, A., and Stemmler, M. (2017) Periodic population codes: From a single circular variable to higher dimensions, multiple nested scales, and conceptual spaces. *Curr. Opin. Neurobiol.*, 46:99–108. CrossRef

Heys, J. G., Rangarajan, K. V., and Dombeck, D. A. (2014) The functional micro-organization of grid cells revealed by cellular-resolution imaging. *Neuron*, 84:1079–1090. CrossRef

Holmgren, C., Harkany, T., Svennenfors, B., and Zilberter, Y. (2003) Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *J. Physiol.*, 551:139–153. CrossRef

Horner, A. J., Bisby, J. A., Zotow, E., Bush, D., and Burgess, N. (2016) Grid-like processing of imagined navigation. *Curr. Biol.*, 26:842–847. CrossRef

Hotelling, H. (1933) Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.*, 24:417–441. CrossRef

Hoydal, O. A., Skytoen, E. R., Moser, M.-B., and Moser, E. I. (2018) Object-vector coding in the medial entorhinal cortex. *bioRxiv*, 0:286286. CrossRef

Huxter, J. R., Senior, T. J., Allen, K., and Csicsvari, J. (2008) Theta phase–specific codes for two-dimensional position, trajectory and heading in the hippocampus. *Nat. Neurosci.*, 11:587–594. CrossRef

Iijima, T., Witter, M. P., Ichikawa, M., Tominaga, T., Kajiwara, R., and Matsumoto, G. (1996) Entorhinal-hippocampal interactions revealed by real-time imaging. *Science*, 272:1176–1179. CrossRef

Insausti, R., Muñoz-López, M., Insausti, A. M., and Artacho-Pérula, E. (2017) The human periallocortex: Layer pattern in presubiculum, parasubiculum and entorhinal cortex. a review. *Front. Neuroanat.*, 11:84. CrossRef

Ismakov, R., Barak, O., Jeffery, K., and Derdikman, D. (2017) Grid cells encode local positional information. *Curr. Biol.*, 27:2337–2343. CrossRef

Jacobs, J., Weidemann, C. T., Miller, J. F., Solway, A., Burke, J. F., Wei, X.-X., Suthana, N., Sperling, M. R., Sharan, A. D., Fried, I., et al. (2013) Direct recordings of grid-like neuronal activity in human spatial navigation. *Nat. Neurosci.*, 16:1188–1190. CrossRef

Jaramillo, J., Schmidt, R., and Kempter, R. (2014) Modeling inheritance of phase precession in the hippocampal formation. *J. Neurosci.*, 34:7715–7731. CrossRef

Jeewajee, A., Barry, C., Douchamps, V., Manson, D., Lever, C., and Burgess, N. (2014) Theta phase precession of grid and place cell firing in open environments. *Phil. Trans. R. Soc. B*, 369:20120532. CrossRef

Julian, J. B., Keinath, A. T., Frazzetta, G., and Epstein, R. A. (2018) Human entorhinal cortex represents visual space using a boundary-anchored grid. *Nat. Neurosci.*, 21:191–194. CrossRef

Jung, M. W., Wiener, S. I., and McNaughton, B. L. (1994) Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat. *J. Neurosci.*, 14:7347–7356.

Keefe, J. O. and Burgess, N. (1996) Geometric determinants of the place fields of hippocampal neurons. *Nature*, 381:425–428. CrossRef

Kempter, R., Gerstner, W., and Van Hemmen, J. L. (1999) Hebbian learning and spiking neurons. *Phys. Rev. E*, 59:4498–4514. CrossRef

Killian, N. J., Jutras, M. J., and Buffalo, E. A. (2012) A map of visual space in the primate entorhinal cortex. *Nature*, 491:761–764. CrossRef

Kim, S., Sapiurka, M., Clark, R. E., and Squire, L. R. (2013) Contrasting effects on path integration after hippocampal damage in humans and rats. *Proc. Nat. Acad. Sci. U.S.A.*, 110:4732–4737. CrossRef

Kim, S. S., Rouault, H., Druckmann, S., and Jayaraman, V. (2017) Ring attractor dynamics in the drosophila central brain. *Science*, 356:849–853. CrossRef

Kitamura, T., Pignatelli, M., Suh, J., Kohara, K., Yoshiki, A., Abe, K., and Tonegawa, S. (2014) Island cells control temporal association memory. *Science*, 343:896–901. CrossRef

Kjelstrup, K. B., Solstad, T., Brun, V. H., Hafting, T., Leutgeb, S., Witter, M. P., Moser, E. I., and Moser, M.-B. (2008) Finite scale of spatial representation in the hippocampus. *Science*, 321:140–143. CrossRef

Klink, R. and Alonso, A. (1997) Morphological characteristics of layer II projection neurons in the rat medial entorhinal cortex. *Hippocampus*, 7:571–583. CrossRef

Kloosterman, F., Van Haeften, T., Witter, M. P., and Lopes da Silva, F. H. (2003) Electrophysiological characterization of interlaminar entorhinal connections: an essential link for re-entrance in the hippocampal–entorhinal system. *Eur. J. Neurosci.*, 18:3037–3052. CrossRef

Ko, H., Cossell, L., Baragli, C., Antolik, J., Clopath, C., Hofer, S. B., and Mrsic-Flogel, T. D. (2013) The emergence of functional microcircuits in visual cortex. *Nature*, 496:96–100. CrossRef

Koenig, J., Linder, A. N., Leutgeb, J. K., and Leutgeb, S. (2011) The spatial periodicity of grid cells is not sustained during reduced theta oscillations. *Science*, 332:592–595. CrossRef

Köhler, C. (1984) Morphological details of the projection from the presubiculum to the entorhinal area as shown with the novel pha-l immunohistochemical tracing method in the rat. *Neurosci. Lett.*, 45:285–290. CrossRef

Kraus, B. J., Brandon, M. P., Robinson, R. J., Connerney, M. A., Hasselmo, M. E., and Eichenbaum, H. (2015) During running in place, grid cells integrate elapsed time and distance run. *Neuron*, 88:578–589. CrossRef

Kropff, E., Carmichael, J. E., Moser, M.-B., and Moser, E. I. (2015) Speed cells in the medial entorhinal cortex. *Nature*, 523:419–424. CrossRef

Kropff, E. and Treves, A. (2008) The emergence of grid cells: Intelligent design or just adaptation? *Hippocampus*, 18:1256–1269. CrossRef

Kruge, I. U., Waaga, T., Wernle, T., Moser, M.-B., and Moser, E. I. (2016) Development of grid cells requires only minimal experience with geometric boundaries. In *Soc. Neurosci. Abstr. (183.12)*.

Krupic, J., Bauza, M., Burton, S., Barry, C., and O'Keefe, J. (2015) Grid cell symmetry is shaped by environmental geometry. *Nature*, 518:232–235. CrossRef

Krupic, J., Bauza, M., Burton, S., and O'Keefe, J. (2016) Framing the grid: effect of boundaries on grid cells and navigation. *J. Physiol.*, 594:6489–6499. CrossRef

Krupic, J., Bauza, M., Burton, S., and O'Keefe, J. (2018) Local transformations of the hippocampal cognitive map. *Science*, 359:1143–1146. CrossRef

Krupic, J., Burgess, N., and O'Keefe, J. (2012) Neural representations of location composed of spatially periodic bands. *Science*, 337:853–857. CrossRef

Kubie, J. L. and Fenton, A. A. (2012) Linear look-ahead in conjunctive cells: an entorhinal mechanism for vector-based navigation. *Front. Neural Circuits*, 6:20. CrossRef

Kumar, S. S. and Buckmaster, P. S. (2006) Hyperexcitability, interneurons, and loss of gabaergic synapses in entorhinal cortex in a model of temporal lobe epilepsy. *J. Neurosci.*, 26:4613–4623. CrossRef

La Camera, G., Rauch, A., Thurbon, D., Lüscher, H.-R., Senn, W., and Fusi, S. (2006) Multiple time scales of temporal response in pyramidal and fast spiking cortical neurons. *J. Neurphysiol.*, 96:3448–3464. CrossRef

Langston, R. F., Ainge, J. A., Couey, J. J., Canto, C. B., Bjerknes, T. L., Witter, M. P., Moser, E. I., and Moser, M.-B. (2010) Development of the spatial representation system in the rat. *Science*, 328:1576–1580. CrossRef

Latuske, P., Kornienko, O., Kohler, L., and Allen, K. (2017) Hippocampal remapping and its entorhinal origin. *Front. Behav. Neurosci.*, 11:253. CrossRef

Lee, S., Hjerling-Leffler, J., Zagha, E., Fishell, G., and Rudy, B. (2010) The largest group of superficial neocortical gabaergic interneurons expresses ionotropic serotonin receptors. *J. Neurosci.*, 30:16796–16808. CrossRef

Lengyel, M., Szatmary, Z., and Erdi, P. (2003) Dynamically detuned oscillations account for the coupled rate and temporal code of place cell firing. *Hippocampus*, 13:700–714. CrossRef

Leutgeb, J. K., Leutgeb, S., Moser, M.-B., and Moser, E. I. (2007) Pattern separation in the dentate gyrus and ca3 of the hippocampus. *Science*, 315:961–966. CrossRef

Leutgeb, S., Leutgeb, J. K., Barnes, C. A., Moser, E. I., McNaughton, B. L., and Moser, M.-B. (2005) Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. *Science*, 309:619–623. CrossRef

Lever, C., Burton, S., Jeewajee, A., O'Keefe, J., and Burgess, N. (2009) Boundary vector cells in the subiculum of the hippocampal formation. *J. Neurosci.*, 29:9771–9777. CrossRef

Lien, A. D. and Scanziani, M. (2013) Tuned thalamic excitation is amplified by visual cortical circuits. *Nat. Neurosci.*, 16:1315–1323. CrossRef

Lisman, J., Buzsáki, G., Eichenbaum, H., Nadel, L., Rangananth, C., and Redish, A. D. (2017) Viewpoints: how the hippocampus contributes to memory, navigation and cognition. *Nat. Neurosci.*, 20:1434–1447. CrossRef

Liu, Y.-H. and Wang, X.-J. (2001) Spike-frequency adaptation of a generalized leaky integrate-and-fire model neuron. *J. Comput. Neurosci.*, 10:25–45.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., and Fry, P. A. (1993) Nonvisual navigation by blind and sighted: assessment of path integration ability. *J. Exp. Psychol.*, 122:73–91. CrossRef

Lorente de Nó, R. (1934) Studies on the structure of the cerebral cortex. II. Continuation of the study of the ammonic system. *J. Psychol. Neurol.*, 46:113–177.

Maaswinkel, H. and Whishaw, I. Q. (1999) Homing with locale, taxon, and dead reckoning strategies by foraging rats: sensory hierarchy in spatial navigation. *Behav. Brain Res.*, 99:143–152. CrossRef

Mallory, C. S., Hardcastle, K., Bant, J. S., and Giocomo, L. M. (2018) Grid scale drives the scale and long-term stability of place maps. *Nat. Neurosci.*, 21:270–282. CrossRef

Manns, I., Mainville, L., and Jones, B. (2001) Evidence for glutamate, in addition to acetylcholine and gaba, neurotransmitter synthesis in basal forebrain neurons projecting to the entorhinal cortex. *Neuroscience*, 107:249–263. CrossRef

Manns, J. R., Howard, M. W., and Eichenbaum, H. (2007) Gradual changes in hippocampal activity support remembering the order of events. *Neuron*, 56:530–540. CrossRef

Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, 275:213–215. CrossRef

Marozzi, E., Ginzberg, L. L., Alenda, A., and Jeffery, K. J. (2015) Purely translational realignment in grid cell firing patterns following nonmetric context change. *Cereb. Cortex*, 25:4619–4627. CrossRef

Mathis, A., Herz, A. V., and Stemmler, M. (2012a) Optimal population codes for space: grid cells outperform place cells. *Neural Comput.*, 24:2280–2317. CrossRef

Mathis, A., Herz, A. V., and Stemmler, M. B. (2012b) Resolution of nested neuronal representations can be exponential in the number of neurons. *Phys. Rev. Lett.*, 109:018103. CrossRef

McFarland, W. L., Teitelbaum, H., and Hedges, E. K. (1975) Relationship between hippocampal theta activity and running speed in the rat. *J. Comp. Physiol. Psychol.*, 88:324–328. CrossRef

McNaughton, B., Chen, L., and Markus, E. (1991) "dead reckoning", landmark learning, and the sense of direction: a neurophysiological and computational hypothesis. *J. Cogn. Neurosci.*, 3:190–202. CrossRef

McNaughton, B. L., Barnes, C. A., Gerrard, J. L., Gothard, K., Jung, M. W., Knierim, J. J., Kudrimoti, H., Qin, Y., Skaggs, W., Suster, M., et al. (1996) Deciphering the hippocampal polyglot: the hippocampus as a path integration system. *J. Exp. Biolog.*, 199:173–185.

McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., and Moser, M.-B. (2006) Path integration and the neural basis of the 'cognitive map'. *Nat. Rev. Neurosci.*, 7:663–678. CrossRef

Meister, M. L. and Buffalo, E. A. (2018) Neurons in primate entorhinal cortex represent gaze position in multiple spatial reference frames. *J. Neurosci.*, 38:2430–2441. CrossRef

Mhatre, H., Gorchetchnikov, A., and Grossberg, S. (2012) Grid cell hexagonal patterns formed by fast self-organized learning within entorhinal cortex. *Hippocampus*, 22:320–334. CrossRef

Miao, C., Cao, Q., Moser, M.-B., and Moser, E. I. (2017) Parvalbumin and somatostatin interneurons control different space-coding networks in the medial entorhinal cortex. *Cell*, 171:507–521. CrossRef

Mishra, R. K., Kim, S., Guzman, S. J., and Jonas, P. (2016) Symmetric spike timing-dependent plasticity at CA3-CA3 synapses optimizes storage and recall in autoassociative networks. *Nat. Commun.*, 7:11552. CrossRef

Mitchell, S., Rawlins, J., Steward, O., and Olton, D. (1982) Medial septal area lesions disrupt theta rhythm and cholinergic staining in medial entorhinal cortex and produce impaired radial arm maze behavior in rats. *J. Neurosci.*, 2:292–302.

Mizuseki, K., Royer, S., Diba, K., and Buzsáki, G. (2012) Activity dynamics and behavioral correlates of CA3 and CA1 hippocampal pyramidal neurons. *Hippocampus*, 22:1659–1680. CrossRef

Monsalve-Mercado, M. M. and Leibold, C. (2017) Hippocampal spike-timing correlations lead to hexagonal grid fields. *Phys. Rev. Lett.*, 119:038101. CrossRef

Morris, R. (2007) *Theories of hippocampal function.* Oxford university press.

Morris, R., Garrud, P., Rawlins, J. a., and O'Keefe, J. (1982) Place navigation impaired in rats with hippocampal lesions. *Nature*, 297:681–683. CrossRef

Moser, E. I., Kropff, E., and Moser, M.-B. (2008) Place cells, grid cells, and the brain's spatial representation system. *Annu. Rev. Neurosci.*, 31:69–89. CrossRef

Moser, E. I., Moser, M.-B., and McNaughton, B. L. (2017) Spatial representation in the hippocampal formation: a history. *Nat. Neurosci.*, 20:1448–1464. CrossRef

Moser, E. I., Roudi, Y., Witter, M. P., Kentros, C., Bonhoeffer, T., and Moser, M.-B. (2014) Grid cells and cortical representation. *Nat. Rev. Neurosci.*, 15:466–481. CrossRef

Müller, M. and Wehner, R. (1988) Path integration in desert ants, cataglyphis fortis. *Proc. Nat. Acad. Sci. U.S.A.*, 85:5287–5290. CrossRef

Muller, R. U. and Kubie, J. L. (1987) The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *J. Neurosci.*, 7:1951–1968.

Murphy, B. K. and Miller, K. D. (2009) Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron*, 61:635–648. CrossRef

Murphy, J. J. (1873) Instinct: A mechanical analogy. *Nature*, 7:483. CrossRef

Murray, J. D. (2002) *Mathematical Biology I: an introduction, Vol. 17 of Interdisciplinary Applied Mathematics* Springer, New York, NY, USA.

Naber, P. A., Caballero-Bleda, M., Jorritsma-Byham, B., and Witter, M. P. (1997) Parallel input to the hippocampal memory system through peri-and postrhinal cortices. *NeuroReport*, 8:2617–2621. CrossRef

Nadel, L. and Moscovitch, M. (1997) Memory consolidation, retrograde amnesia and the hippocampal complex. *Curr. Opin. Neurobiol.*, 7:217–227. CrossRef

Nau, M., Schröder, T. N., Bellmund, J. L., and Doeller, C. F. (2018) Hexadirectional coding of visual space in human entorhinal cortex. *Nat. Neurosci.*, 21:188–190. CrossRef

Naumann, R., Preston-Ferrer, P., Brecht, M., and Burgalossi, A. (2018) Structural modularity and grid activity in the medial entorhinal cortex. *J. Neurophysiol.* CrossRef In press

Naumann, R. K., Ray, S., Prokop, S., Las, L., Heppner, F. L., and Brecht, M. (2016) Conserved size and periodicity of pyramidal patches in layer 2 of medial/caudal entorhinal cortex. *J. Comp. Neurol.*, 524:783–806. CrossRef

Navratilova, Z., Giocomo, L. M., Fellous, J.-M., Hasselmo, M. E., and McNaughton, B. L. (2012) Phase precession and variable spatial scaling in a periodic attractor map model of medial entorhinal grid cells with realistic after-spike dynamics. *Hippocampus*, 22:772–789. CrossRef

Navratilova, Z., Godfrey, K. B., and McNaughton, B. L. (2015) Grids from bands, or bands from grids? an examination of the effects of single unit contamination on grid cell firing fields. *J. Neurophysiol.*, 115:992–1002. CrossRef

Newman, E. L. and Hasselmo, M. E. (2014) Grid cell firing properties vary as a function of theta phase locking preferences in the rat medial entorhinal cortex. *Front. Syst. Neurosci.*, 8:193. CrossRef

Nicoll, R. and Alger, B. (1981) Synaptic excitation may activate a calcium-dependent potassium conductance in hippocampal pyramidal cells. *Science*, 212:957–959. CrossRef

Nolan, M. F., Dudman, J. T., Dodson, P. D., and Santoro, B. (2007) HCN1 channels control resting and active integrative properties of stellate cells from layer II of the entorhinal cortex. *J. Neurosci.*, 27:12440–12451. CrossRef

Oja, E. (1982) Simplified neuron model as a principal component analyzer. *J. Math. Biol.*, 15:267–273. CrossRef

O'Keefe, J. (1976) Place units in the hippocampus of the freely moving rat. *Exp. Neurol.*, 51:78–109. CrossRef

O'Keefe, J. and Burgess, N. (2005) Dual phase and rate coding in hippocampal place cells: theoretical significance and relationship to entorhinal grid cells. *Hippocampus*, 15:853–866. CrossRef

O'Keefe, J. and Dostrovsky, J. (1971) The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain Res.*, 34:171–175. CrossRef

O'Keefe, J. and Nadel, L. (1978) *The hippocampus as a cognitive map* Oxford: Clarendon Press. CrossRef

O'Keefe, J. and Recce, M. L. (1993) Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, 3:317–330. CrossRef

Ólafsdóttir, H. F., Bush, D., and Barry, C. (2018) The role of hippocampal replay in memory and planning. *Curr. Biol.*, 28:R37–R50. CrossRef

Ólafsdóttir, H. F., Carpenter, F., and Barry, C. (2016) Coordinated grid and place cell replay during rest. *Nat. Neurosci.*, 19:792–794. CrossRef

Olton, D. S., Becker, J. T., and Handelmann, G. E. (1979) Hippocampus, space, and memory. *Behav. Brain Sci.*, 2:313–322. CrossRef

O'Mara, S. M., Commins, S., Anderson, M., and Gigg, J. (2001) The subiculum: a review of form, physiology and function. *Prog. Neurobiol.*, 64:129–155. CrossRef

Omer, D. B., Maimon, S. R., Las, L., and Ulanovsky, N. (2018) Social place-cells in the bat hippocampus. *Science*, 359:218–224. CrossRef

O'Neill, J., Boccara, C., Stella, F., Schoenenberger, P., and Csicsvari, J. (2017) Superficial layers of the medial entorhinal cortex replay independently of the hippocampus. *Science*, 355:184–188. CrossRef

O'Brien, R. J., Kamboj, S., Ehlers, M. D., Rosen, K. R., Fischbach, G. D., and Huganir, R. L. (1998) Activity-dependent modulation of synaptic AMPA receptor accumulation. *Neuron*, 21:1067–1078. CrossRef

Pastoll, H., Ramsden, H., and Nolan, M. F. (2012) Intrinsic electrophysiological properties of entorhinal cortex stellate cells and their contribution to grid cell firing fields. *Front. Neural Circuits*, 6:17. CrossRef

Pastoll, H., Solanka, L., van Rossum, M. C., and Nolan, M. F. (2013) Feedback inhibition enables theta-nested gamma oscillations and grid firing fields. *Neuron*, 77:141–154. CrossRef

Pérez-Escobar, J. A., Kornienko, O., Latuske, P., Kohler, L., and Allen, K. (2016) Visual landmarks sharpen grid cell metric and confer context specificity to neurons of the medial entorhinal cortex. *eLife*, 5:e16937. CrossRef

Perin, R., Berger, T. K., and Markram, H. (2011) A synaptic organizing principle for cortical neuronal groups. *Proc. Nat. Acad. Sci. U.S.A.*, 108:5419–5424. CrossRef

Pilly, P. K. and Grossberg, S. (2012) How do spatial learning and memory occur in the brain? coordinated learning of entorhinal grid cells and hippocampal place cells. *J. Cogn. Neurosci.*, 24:1031–1054. CrossRef

Pilly, P. K. and Grossberg, S. (2013) Spiking neurons in a hierarchical self-organizing map model can learn to develop spatial and temporal properties of entorhinal grid cells and hippocampal place cells. *PLOS One*, 8:e60599. CrossRef

Pouget, A., Dayan, P., and Zemel, R. (2000) Information processing with population codes. *Nat. Rev. Neurosci.*, 1:125–132. CrossRef

Preston-Ferrer, P., Coletta, S., Frey, M., and Burgalossi, A. (2016) Anatomical organization of presubicular head-direction circuits. *eLife*, 5:e14592. CrossRef

Quirk, G. J., Muller, R. U., Kubie, J. L., and Ranck, J. (1992) The positional firing properties of medial entorhinal neurons: description and comparison with hippocampal place cells. *J. Neurosci.*, 12:1945–1963.

Rank, J. B. (1984) Head-direction cells in the deep layers of the dorsal presubiculum in freely moving rats. In *Soc. Neurosci. Abstr. (10.599)*.

Rawlins, J. and Olton, D. (1982) The septo-hippocampal system and cognitive mapping. *Behav. Brain Res.*, 5:331–358. CrossRef

Ray, S., Burgalossi, A., Brecht, M., and Naumann, R. K. (2017) Complementary modular microcircuits of the rat medial entorhinal cortex. *Front. Syst. Neurosci.*, 11:20. CrossRef

Ray, S., Naumann, R., Burgalossi, A., Tang, Q., Schmidt, H., and Brecht, M. (2014) Grid-layout and theta-modulation of layer 2 pyramidal neurons in medial entorhinal cortex. *Science*, 343:891–896. CrossRef

Redish, A. D. (1999) *Beyond the cognitive map: from place cells to episodic memory* MIT Press.

Reifenstein, E. T., Kempter, R., Schreiber, S., Stemmler, M. B., and Herz, A. V. (2012) Grid cells in rat entorhinal cortex encode physical space with independent firing fields and phase precession at the single-trial level. *Proc. Nat. Acad. Sci. U.S.A.*, 109:6301–6306. CrossRef

Reifenstein, E. T., Stemmler, M., Herz, A. V., Kempter, R., and Schreiber, S. (2014) Movement dependence and layer specificity of entorhinal phase precession in two-dimensional environments. *PLOS One*, 9:e100638. CrossRef

Remme, M. W., Lengyel, M., and Gutkin, B. S. (2010) Democracy-independence trade-off in oscillating dendrites and its implications for grid cells. *Neuron*, 66:429–437. CrossRef

Rich, P. D., Liaw, H.-P., and Lee, A. K. (2014) Large environments reveal the statistical structure governing hippocampal representations. *Science*, 345:814–817. CrossRef

Richter, H., Heinemann, U., and Eder, C. (2000) Hyperpolarization-activated cation currents in stellate and pyramidal neurons of rat entorhinal cortex. *Neurosci. Lett.*, 281:33–36. CrossRef

Robinson, D. A. (1989) Integrating with neurons. *Annu. Rev. Neurosci.*, 12:33–45. CrossRef

Robinson, R. B. and Siegelbaum, S. A. (2003) Hyperpolarization-activated cation currents: from molecules to physiological function. *Annu. Rev. Physiol.*, 65:453–480. CrossRef

Rolls, E. T., Stringer, S. M., and Elliot, T. (2006) Entorhinal cortex grid cells can map to hippocampal place cells by competitive learning. *Netw. Comput. Neural Syst.*, 17:447–465. CrossRef

Rolls, E. T., Treves, A., and Rolls, E. T. (1998) *Neural networks and brain function*, volume 572 Oxford university press Oxford. CrossRef

Ronacher, B. and Wehner, R. (1995) Desert ants cataglyphis fortis use self-induced optic flow to measure distances travelled. *J. Comp. Physiol. A*, 177:21–27. CrossRef

Rowland, D. C., Roudi, Y., Moser, M.-B., and Moser, E. I. (2016) Ten years of grid cells. *Annu. Rev. Neurosci.*, 39:19–40. CrossRef

Roy, D. S. and Tonegawa, S. (2017) Manipulating memory in space and time. *Curr. Opin. Behav. Sci.*, 17:1–6. CrossRef

Rudy, B., Fishell, G., Lee, S., and Hjerling-Leffler, J. (2011) Three groups of interneurons account for nearly 100% of neocortical gabaergic neurons. *Dev. Neurobiol.*, 71:45–61. CrossRef

Samsonovich, A. and McNaughton, B. L. (1997) Path integration and cognitive mapping in a continuous attractor neural network model. *J. Neurosci.*, 17:5900–5920.

Sarel, A., Finkelstein, A., Las, L., and Ulanovsky, N. (2017) Vectorial representation of spatial goals in the hippocampus of bats. *Science*, 355:176–180. CrossRef

Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B. L., Witter, M. P., Moser, M.-B., and Moser, E. I. (2006) Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science*, 312:758–762. CrossRef

Savelli, F., Yoganarasimha, D., and Knierim, J. J. (2008) Influence of boundary removal on the spatial representations of the medial entorhinal cortex. *Hippocampus*, 18:1270–1282. CrossRef

Schlesiger, M. I., Boublil, B. L., Hales, J. B., Leutgeb, J. K., and Leutgeb, S. (2018) Hippocampal global remapping can occur without input from the medial entorhinal cortex. *Cell Rep.*, 22:3152–3159. CrossRef

Schmidt, H., Gour, A., Straehle, J., Boergens, K. M., Brecht, M., and Helmstaedter, M. (2017) Axonal synapse sorting in medial entorhinal cortex. *Nature*, 549:469–475. CrossRef

Schmidt, R., Diba, K., Leibold, C., Schmitz, D., Buzsáki, G., and Kempter, R. (2009) Single-trial phase precession in the hippocampus. *J. Neurosci.*, 29:13232–13241. CrossRef

Schmidt-Hieber, C. and Häusser, M. (2013) Cellular mechanisms of spatial navigation in the medial entorhinal cortex. *Nat. Neurosci.*, 16:325–331. CrossRef

Schreiber, S., Erchova, I., Heinemann, U., and Herz, A. V. (2004) Subthreshold resonance explains the frequency-dependent integration of periodic as well as random stimuli in the entorhinal cortex. *J. Neurophysiol.*, 92:408–415. CrossRef

Scoville, W. B. and Milner, B. (1957) Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry*, 20:11. CrossRef

Seelig, J. D. and Jayaraman, V. (2015) Neural dynamics for landmark orientation and angular path integration. *Nature*, 521:186–191. CrossRef

Seung, H. S. (1998) Continuous attractors and oculomotor control. *Neural Networks*, 11:1253–1258. CrossRef

Shay, C. F., Ferrante, M., Chapman, G. W., and Hasselmo, M. E. (2016) Rebound spiking in layer ii medial entorhinal cortex stellate cells: Possible mechanism of grid cell function. *Neurobiol. Learn. Mem.*, 129:83–98. CrossRef

Si, B., Kropff, E., and Treves, A. (2012) Grid alignment in entorhinal cortex. *Biol. Cybern.*, 106:483–506. CrossRef

Si, B. and Treves, A. (2013) A model for the differentiation between grid and conjunctive units in medial entorhinal cortex. *Hippocampus*, 23:1410–1424. CrossRef

Skaggs, W. E., Knierim, J. J., Kudrimoti, H. S., and McNaughton, B. L. (1995) A model of the neural basis of the rat's sense of direction. In *Adv. Neural Inf. Process. Syst.*, pages 173–180.

Skaggs, W. E. and McNaughton, B. L. (1996) Theta phase precession in hippocampal neural populations and the compression of temporal sequences. *Hippocampus*, 6:149–172.

Sławińska, U. and Kasicki, S. (1998) The frequency of rat's hippocampal theta rhythm is related to the speed of locomotion. *Brain Res.*, 796:327–331. CrossRef

Slomianka, L. and Geneser, F. (1991) Distribution of acetylcholinesterase in the hippocampal region of the mouse: I. entorhinal area, parasubiculum, retrosplenial area, and presubiculum. *J. Comp. Neurol.*, 303:339–354. CrossRef

Slomianka, L. and Geneser, F. A. (1997) Postnatal development of zinc-containing cells and neuropil in the hippocampal region of the mouse. *Hippocampus*, 7:321–340. CrossRef

Solger, J., Wozny, C., Manahan-Vaughan, D., and Behr, J. (2004) Distinct mechanisms of bidirectional activity-dependent synaptic plasticity in superficial and deep layers of rat entorhinal cortex. *Eur. J. Neurosci.*, 19:2003–2007. CrossRef

Solstad, T., Boccara, C. N., Kropff, E., Moser, M.-B., and Moser, E. I. (2008) Representation of geometric borders in the entorhinal cortex. *Science*, 322:1865–1868. CrossRef

Solstad, T., Moser, E. I., and Einevoll, G. T. (2006) From grid cells to place cells: a mathematical model. *Hippocampus*, 16:1026–1031. CrossRef

Somers, D. C., Nelson, S. B., and Sur, M. (1995) An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci.*, 15:5448–5465.

Song, S., Miller, K. D., and Abbott, L. F. (2000) Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nat. Neurosci.*, 3:919–926. CrossRef

Spencer, S. S. and Spencer, D. D. (1994) Entorhinal-hippocampal interactions in medial temporal lobe epilepsy. *Epilepsia*, 35:721–727. CrossRef

Squire, L. R. (1992) Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.*, 99:195–231. CrossRef

Squire, L. R. and Zola-Morgan, S. (1991) The medial temporal lobe memory system. *Science*, 253:1380–1386. CrossRef

Sreenivasan, S. and Fiete, I. (2011) Grid cells generate an analog error-correcting code for singularly precise neural computation. *Nat. Neurosci.*, 14:1330–1337. CrossRef

Stangl, M., Achtzehn, J., Huber, K., Dietrich, C., Tempelmann, C., and Wolbers, T. (2018) Compromised grid-cell-like representations in old age as a key mechanism to explain age-related navigational deficits. *Curr. Biol.*, 28:1108–1115. CrossRef

Stella, F., Si, B., Kropff, E., and Treves, A. (2013) Grid cells on the ball. *J. Stat. Mech. Theor. Exp.*, 2013:P03013. CrossRef

Stemmler, M., Mathis, A., and Herz, A. V. (2015) Connecting multiple spatial scales to decode the population activity of grid cells. *Sci. Adv.*, 1:e1500816. CrossRef

Stensola, H., Stensola, T., Solstad, T., Frøland, K., Moser, M.-B., and Moser, E. I. (2012) The entorhinal grid map is discretized. *Nature*, 492:72–78. CrossRef

Stensola, T., Stensola, H., Moser, M.-B., and Moser, E. I. (2015) Shearing-induced asymmetry in entorhinal grid cells. *Nature*, 518:207–212. CrossRef

Stepanyants, A. and Chklovskii, D. B. (2005) Neurogeometry and potential synaptic connectivity. *Trends Neurosci.*, 28:387–394. CrossRef

Stepanyuk, A. (2015) Self-organization of grid fields under supervision of place cells in a neuron model with associative plasticity. *Biol. Insp. Cogn. Archit.*, 13:48–62. CrossRef

Stimberg, M., Goodman, D. F., Benichoux, V., and Brette, R. (2014) Equation-oriented specification of neural models for simulations. *Front. Neuroinform.*, 8:6. CrossRef

Stringer, S., Rolls, E., Trappenberg, T., and De Araujo, I. (2002a) Self-organizing continuous attractor networks and path integration: two-dimensional models of place cells. *Netw. Comput. Neural Syst.*, 13:429–446. CrossRef

Stringer, S., Trappenberg, T., Rolls, E., and De Araujo, I. (2002b) Self-organizing continuous attractor networks and path integration: one-dimensional models of head direction cells. *Netw. Comput. Neural Syst.*, 13:217–242. CrossRef

Strösslin, T., Sheynikhovich, D., Chavarriaga, R., and Gerstner, W. (2005) Robust self-localisation and navigation based on hippocampal place cells. *Neural Networks*, 18:1125–1140. CrossRef

Suarez, H., Koch, C., and Douglas, R. (1995) Modeling direction selectivity of simple cells in striate visual cortex within the framework of the canonical microcircuit. *J. Neurosci.*, 15:6700–6719.

Sun, C., Kitamura, T., Yamamoto, J., Martin, J., Pignatelli, M., Kitch, L. J., Schnitzer, M. J., and Tonegawa, S. (2015) Distinct speed dependence of entorhinal island and ocean cells, including respective grid cells. *Proc. Nat. Acad. Sci. U.S.A.*, 112:9466–9471. CrossRef

Sürmeli, G., Marcu, D. C., McClure, C., Garden, D. L., Pastoll, H., and Nolan, M. F. (2015) Molecularly defined circuitry reveals input-output segregation in deep layers of the medial entorhinal cortex. *Neuron*, 88:1040–1053. CrossRef

Tamamaki, N. and Nojyo, Y. (1995) Preservation of topography in the connections between the subiculum, field CA1, and the entorhinal cortex in rats. *J. Comp. Neurol.*, 353:379–390. CrossRef

Tang, Q., Burgalossi, A., Ebbesen, C. L., Ray, S., Naumann, R., Schmidt, H., Spicher, D., and Brecht, M. (2014) Pyramidal and stellate cell specificity of grid and border representations in layer 2 of medial entorhinal cortex. *Neuron*, 84:1191–1197. CrossRef

Tang, Q., Burgalossi, A., Ebbesen, C. L., Sanguinetti-Scheck, J. I., Schmidt, H., Tukker, J. J., Naumann, R., Ray, S., Preston-Ferrer, P., Schmitz, D., and Brecht, M. (2016) Functional architecture of the rat parasubiculum. *J. Neurosci.*, 36:2289–2301. CrossRef

Taube, J. S. (1995a) Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *J. Neurosci.*, 15:70–86.

Taube, J. S. (1995b) Place cells recorded in the parasubiculum of freely moving rats. *Hippocampus*, 5:569–583. CrossRef

Taube, J. S., Muller, R. U., and Ranck, J. B. (1990) Head-direction cells recorded from the postsubiculum in freely moving rats. i. description and quantitative analysis. *J. Neurosci.*, 10:420–435.

Tavares, R. M., Mendelsohn, A., Grossman, Y., Williams, C. H., Shapiro, M., Trope, Y., and Schiller, D. (2015) A map for social navigation in the human brain. *Neuron*, 87:231–243. CrossRef

Taylor, J. (1999) Neural 'bubble' dynamics in two dimensions: foundations. *Biol. Cybern.*, 80:393–409. CrossRef

Taylor, K. (1978) Range of movement and activity of common rats (rattus norvegicus) on agricultural land. *J. Appl. Ecol.*, 15:663–677. CrossRef

Tennant, S. A., Fischer, L., Garden, D. L., Gerlei, K. Z., Martinez-Gonzalez, C., McClure, C., Wood, E. R., and Nolan, M. F. (2018) Stellate cells in the medial entorhinal cortex are required for spatial learning. *Cell Rep.*, 22:1313–1324. CrossRef

Thue, A. (1910) *Über die dichteste Zusammenstellung von kongruenten Kreisen in einer Ebene* J. Dybwad.

Thurley, K., Hellmundt, F., and Leibold, C. (2013) Phase precession of grid cells in a network model without external pacemaker. *Hippocampus*, 23:786–796. CrossRef

Tocker, G., Barak, O., and Derdikman, D. (2015) Grid cells correlation structure suggests organized feedforward projections into superficial layers of the medial entorhinal cortex. *Hippocampus*, 25:1599–1613. CrossRef

Tolman, E. C. (1948) Cognitive maps in rats and men. *Psycholog. Rev.*, 55:189–208. CrossRef

Tulving, E. and Markowitsch, H. J. (1998) Episodic and declarative memory: role of the hippocampus. *Hippocampus*, 8:198–204. CrossRef

Turing, A. M. (1952) The chemical basis of morphogenesis. *Phil. Trans. R. Soc. B*, 237:37–72. CrossRef

Turner-Evans, D., Wegener, S., Rouault, H., Franconville, R., Wolff, T., Seelig, J. D., Druckmann, S., and Jayaraman, V. (2017) Angular velocity integration in a fly heading circuit. *eLife*, 6:e23496. CrossRef

Turrigiano, G. G., Leslie, K. R., Desai, N. S., Rutherford, L. C., and Nelson, S. B. (1998) Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature*, 391:892–896. CrossRef

Turrigiano, G. G. and Nelson, S. B. (2004) Homeostatic plasticity in the developing nervous system. *Nat. Rev. Neurosci.*, 5:97–107. CrossRef

Urdapilleta, E., Si, B., and Treves, A. (2017) Self-organization of modular activity of grid cells. *Hippocampus*, 27:1204–1213. CrossRef

Van Cauter, T., Poucet, B., and Save, E. (2008) Unstable CA1 place cell representation in rats with entorhinal cortex lesions. *Eur. J. Neurosci.*, 27:1933–1946. CrossRef

Van der Linden, S. and da Silva, F. H. L. (1998) Comparison of the electrophysiology and morphology of layers III and II neurons of the rat medial entorhinal cortex in vitro. *Eur. J. Neurosci.*, 10:1479–1489. CrossRef

Van Groen, T. and Wyss, J. M. (1990) The connections of presubiculum and parasubiculum in the rat. *Brain Res.*, 518:227–243. CrossRef

Van Haeften, T., Baks-te Bulte, L., Goede, P. H., Wouterlood, F. G., and Witter, M. P. (2003) Morphological and numerical analysis of synaptic interactions between neurons in deep and superficial layers of the entorhinal cortex of the rat. *Hippocampus*, 13:943–952. CrossRef

Van Hoesen, G. W., Hyman, B. T., and Damasio, A. R. (1991) Entorhinal cortex pathology in Alzheimer's disease. *Hippocampus*, 1:1–8.

Van Strien, N. M., Cappaert, N., and Witter, M. P. (2009) The anatomy of memory: an interactive overview of the parahippocampal-hippocampal network. *Nat. Rev. Neurosci.*, 10:272–282. CrossRef

Vanderwolf, C. (1969) Hippocampal electrical activity and voluntary movement in the rat. *Electroencephalogr. Clin. Neurophysiol.*, 26:407–418.

Varga, C., Lee, S. Y., and Soltesz, I. (2010) Target-selective GABAergic control of entorhinal cortex output. *Nat. Neurosci.*, 13:822–824. CrossRef

Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., and Mishkin, M. (1997) Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, 277:376–380. CrossRef

Vucurovic, K., Gallopin, T., Ferezou, I., Rancillac, A., Chameau, P., van Hooft, J. A., Geoffroy, H., Monyer, H., Rossier, J., and Vitalis, T. (2010) Serotonin 3a receptor subtype as an early and protracted marker of cortical interneuron subpopulations. *Cereb. Cortex*, 20:2333–2347. CrossRef

Wallace, A. R. (1873) Inherited feeling. *Nature*, 7:303. CrossRef

Wallraff, H. G. (2005) *Avian navigation: pigeon homing as a paradigm* Springer Science & Business Media.

Weber, S. N. and Sprekeler, H. (2018) Learning place cells, grid cells and invariances with excitatory and inhibitory plasticity. *eLife*, 7:e34560. CrossRef

Wehner, R. d. (1976) Polarized-light navigation by insects. *Sci. Am.*, 235:106–115. CrossRef

Wei, X.-X., Prentice, J., and Balasubramanian, V. (2015) A principle of economy predicts the functional architecture of grid cells. *eLife*, 4:e08362. CrossRef

Welday, A. C., Shlifer, I. G., Bloom, M. L., Zhang, K., and Blair, H. T. (2011) Cosine directional tuning of theta cell burst frequencies: evidence for spatial coding by oscillatory interference. *J. Neurosci.*, 31:16157–16176. CrossRef

Welinder, P. E., Burak, Y., and Fiete, I. R. (2008) Grid cells: the position code, neural network models of activity, and the problem of learning. *Hippocampus*, 18:1283–1300. CrossRef

Wernle, T., Waaga, T., Mørreaunet, M., Treves, A., Moser, M.-B., and Moser, E. I. (2018) Integration of grid maps in merged environments. *Nat. Neurosci.*, 21:92–101. CrossRef

Widloski, J. and Fiete, I. R. (2014) A model of grid cell development through spatial exploration and spike time-dependent plasticity. *Neuron*, 83:481–495. CrossRef

Wills, T. J., Barry, C., and Cacucci, F. (2012) The abrupt development of adult-like grid cell firing in the medial entorhinal cortex. *Front. Neural Circuits*, 6:21. CrossRef

Wills, T. J., Cacucci, F., Burgess, N., and O'Keefe, J. (2010) Development of the hippocampal cognitive map in preweanling rats. *Science*, 328:1573–1576. CrossRef

Wills, T. J., Lever, C., Cacucci, F., Burgess, N., and O'keefe, J. (2005) Attractor dynamics in the hippocampal representation of the local environment. *Science*, 308:873–876. CrossRef

Wilming, N., König, P., König, S., and Buffalo, E. A. (2018) Entorhinal cortex receptive fields are modulated by spatial attention, even without movement. *eLife*, 7:e31745. CrossRef

Wilson, H. R. and Cowan, J. D. (1972) Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.*, 12:1–24. CrossRef

Wilson, H. R. and Cowan, J. D. (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13:55–80. CrossRef

Winson, J. (1974) Patterns of hippocampal theta rhythm in the freely moving rat. *Electroencephalogr. Clin. Neurophysiol.*, 36:291–301. CrossRef

Winter, S. S., Clark, B. J., and Taube, J. S. (2015) Disruption of the head direction cell network impairs the parahippocampal grid cell signal. *Science*, 347:870–874. CrossRef

Winterer, J., Maier, N., Wozny, C., Beed, P., Breustedt, J., Evangelista, R., Peng, Y., D'Albis, T., Kempter, R., and Schmitz, D. (2017) Excitatory microcircuits within superficial layers of the medial entorhinal cortex. *Cell Rep.*, 19:1110–1116. CrossRef

Witter, M. P., Doan, T. P., Jacobsen, B., Nilssen, E. S., and Ohara, S. (2017) Architecture of the entorhinal cortex a review of entorhinal anatomy in rodents with some comparative notes. *Front. Syst. Neurosci.*, 11:46. CrossRef

Witter, M. P., Naber, P. A., van Haeften, T., Machielsen, W. C., Rombouts, S. A., Barkhof, F., Scheltens, P., and Lopes da Silva, F. H. (2000) Cortico-hippocampal communication by way of parallel parahippocampal-subicular pathways. *Hippocampus*, 10:398–410. CrossRef

Wittlinger, M., Wehner, R., and Wolf, H. (2006) The ant odometer: stepping on stilts and stumps. *Science*, 312:1965–1967. CrossRef

Wu, S., Hamaguchi, K., and Amari, S.-i. (2008) Dynamics and computation of continuous attractors. *Neural Comput.*, 20:994–1025. CrossRef

Yang, S., Lee, D., Chung, C., Cheong, M., Lee, C.-J., and Jung, M. (2004) Long-term synaptic plasticity in deep layer-originated associational projections to superficial layers of rat entorhinal cortex. *Neuroscience*, 127:805–812. CrossRef

Yartsev, M. M., Witter, M. P., and Ulanovsky, N. (2011) Grid cells without theta oscillations in the entorhinal cortex of bats. *Nature*, 479:103–107. CrossRef

Yoon, K., Buice, M. A., Barry, C., Hayman, R., Burgess, N., and Fiete, I. R. (2013) Specific evidence of low-dimensional continuous attractor dynamics in grid cells. *Nat. Neurosci.*, 16:1077–1084. CrossRef

Yoshida, M., Jochems, A., and Hasselmo, M. E. (2013) Comparison of properties of medial entorhinal cortex layer II neurons in two anatomical dimensions with and without cholinergic activation. *PLOS One*, 8:e73904. CrossRef

Yun, S. H., Mook-Jung, I., and Jung, M. W. (2002) Variation in effective stimulus patterns for induction of long-term potentiation across different layers of rat entorhinal cortex. *J. Neurosci.*, 22:RC214.

Zhang, K. (1996) Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J. Neurosci.*, 16:2112–2126.

Zhang, S.-J., Ye, J., Miao, C., Tsao, A., Cerniauskas, I., Ledergerber, D., Moser, M.-B., and Moser, E. I. (2013) Optogenetic dissection of entorhinal-hippocampal functional connectivity. *Science*, 340:1232627. CrossRef

Zhou, Y.-D., Acker, C. D., Netoff, T. I., Sen, K., and White, J. A. (2005) Increasing Ca2+ transients by broadening postsynaptic action potentials enhances timing-dependent synaptic depression. *Proc. Nat. Acad. Sci. U.S.A.*, 102:19121–19125. CrossRef

Zilli, E. A. and Hasselmo, M. E. (2010) Coupled noisy spiking neurons as velocity-controlled oscillators in a model of grid cell spatial firing. *J. Neurosci.*, 30:13850–13860. CrossRef

Zilli, E. A., Yoshida, M., Tahvildari, B., Giocomo, L. M., and Hasselmo, M. E. (2009) Evaluation of the oscillatory interference model of grid cell firing through analysis and measured period variance of some biological oscillators. *PLOS Comput. Biol.*, 5:e1000573. CrossRef

Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, den

Tiziano D'Albis